

Based upon the consensus sequences obtained as described above, oligonucleotides were then synthesized and used to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for a PRO polypeptide. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

EXAMPLE 2: Isolation of cDNA clones by Amylase Screening

1. Preparation of oligo dT primed cDNA library

mRNA was isolated from a human tissue of interest using reagents and protocols from Invitrogen, San Diego, CA (Fast Track 2). This RNA was used to generate an oligo dT primed cDNA library in the vector pRK5D using reagents and protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System). In this procedure, the double stranded cDNA was sized to greater than 1000 bp and the SalI/NotI linked cDNA was cloned into XhoI/NotI cleaved vector. pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites.

2. Preparation of random primed cDNA library

A secondary cDNA library was generated in order to preferentially represent the 5' ends of the primary cDNA clones. Sp6 RNA was generated from the primary library (described above), and this RNA was used to generate a random primed cDNA library in the vector pSST-AMY.0 using reagents and protocols from Life Technologies (Super Script Plasmid System, referenced above). In this procedure the double stranded cDNA was sized to 500-1000 bp, linked with blunt to NotI adaptors, cleaved with SfiI, and cloned into SfiI/NotI cleaved vector. pSST-AMY.0 is a cloning vector that has a yeast alcohol dehydrogenase promoter preceding the cDNA cloning sites and the mouse amylase sequence (the mature sequence without the secretion signal) followed by the yeast alcohol dehydrogenase terminator, after the cloning sites. Thus, cDNAs cloned into this vector that are fused in frame with amylase sequence will lead to the secretion of amylase from appropriately transfected yeast colonies.

3. Transformation and Detection

DNA from the library described in paragraph 2 above was chilled on ice to which was added electrocompetent DH10B bacteria (Life Technologies, 20 ml). The bacteria and vector mixture was then electroporated as recommended by the manufacturer. Subsequently, SOC media (Life Technologies, 1 ml) was added and the mixture was incubated at 37°C for 30 minutes. The transformants were then plated onto 20 standard 150 mm LB plates containing ampicillin and incubated for 16 hours (37°C). Positive colonies were scraped off the plates and the DNA was isolated from the bacterial pellet using standard protocols, e.g. CsCl-gradient. The purified DNA was then carried on to the yeast protocols below.

The yeast methods were divided into three categories: (1) Transformation of yeast with the plasmid/cDNA combined vector; (2) Detection and isolation of yeast clones secreting amylase; and (3) PCR amplification of the insert directly from the yeast colony and purification of the DNA for sequencing and further analysis.

The yeast strain used was HD56-5A (ATCC-90785). This strain has the following genotype: MAT alpha, ura3-52, leu2-3, leu2-112, his3-11, his3-15, MAL⁺, SUC⁺, GAL⁺. Preferably, yeast mutants can be employed that have deficient post-translational pathways. Such mutants may have translocation deficient alleles in *sec71*, *sec72*, *sec62*, with truncated *sec71* being most preferred. Alternatively, antagonists (including antisense nucleotides and/or ligands) which interfere with the normal operation of these genes, other proteins implicated in this post translation pathway (e.g., SEC61p, SEC72p, SEC62p, SEC63p, TDJ1p or SSA1p-4p) or the complex formation of these proteins may also be preferably employed in combination with the amylase-expressing yeast.

Transformation was performed based on the protocol outlined by Gietz et al., Nucl. Acid. Res., 20:1425 (1992). Transformed cells were then inoculated from agar into YEPD complex media broth (100 ml) and grown overnight at 30°C. The YEPD broth was prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 207 (1994). The overnight culture was then diluted to about 2×10^6 cells/ml (approx. OD₆₀₀=0.1) into fresh YEPD broth (500 ml) and regrown to 1×10^7 cells/ml (approx. OD₆₀₀=0.4-0.5).

The cells were then harvested and prepared for transformation by transfer into GS3 rotor bottles in a Sorval GS3 rotor at 5,000 rpm for 5 minutes, the supernatant discarded, and then resuspended into sterile water, and centrifuged again in 50 ml falcon tubes at 3,500 rpm in a Beckman GS-6KR centrifuge. The supernatant was discarded and the cells were subsequently washed with LiAc/TE (10 ml, 10 mM Tris-HCl, 1 mM EDTA pH 7.5, 100 mM Li₂OOCCH₃), and resuspended into LiAc/TE (2.5 ml).

Transformation took place by mixing the prepared cells (100 µl) with freshly denatured single stranded salmon testes DNA (Lofstrand Labs, Gaithersburg, MD) and transforming DNA (1 µg, vol. < 10 µl) in microfuge tubes. The mixture was mixed briefly by vortexing, then 40% PEG/TE (600 µl, 40% polyethylene glycol-4000, 10 mM Tris-HCl, 1 mM EDTA, 100 mM Li₂OOCCH₃, pH 7.5) was added. This mixture was gently mixed and incubated at 30°C while agitating for 30 minutes. The cells were then heat shocked at 42°C for 15 minutes, and the reaction vessel centrifuged in a microfuge at 12,000 rpm for 5-10 seconds, decanted and resuspended into TE (500 µl, 10 mM Tris-HCl, 1 mM EDTA pH 7.5) followed by recentrifugation. The

cells were then diluted into TE (1 ml) and aliquots (200 μ l) were spread onto the selective media previously prepared in 150 mm growth plates (VWR).

Alternatively, instead of multiple small reactions, the transformation was performed using a single, large scale reaction, wherein reagent amounts were scaled up accordingly.

The selective media used was a synthetic complete dextrose agar lacking uracil (SCD-Ura) prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 208-210 (1994). Transformants were grown at 30°C for 2-3 days.

The detection of colonies secreting amylase was performed by including red starch in the selective growth media. Starch was coupled to the red dye (Reactive Red-120, Sigma) as per the procedure described by Biely et al., Anal. Biochem., 172:176-179 (1988). The coupled starch was incorporated into the SCD-Ura agar plates at a final concentration of 0.15% (w/v), and was buffered with potassium phosphate to a pH of 7.0 (50-100 mM final concentration).

The positive colonies were picked and streaked across fresh selective media (onto 150 mm plates) in order to obtain well isolated and identifiable single colonies. Well isolated single colonies positive for amylase secretion were detected by direct incorporation of red starch into buffered SCD-Ura agar. Positive colonies were determined by their ability to break down starch resulting in a clear halo around the positive colony visualized directly.

4. Isolation of DNA by PCR Amplification

When a positive colony was isolated, a portion of it was picked by a toothpick and diluted into sterile water (30 μ l) in a 96 well plate. At this time, the positive colonies were either frozen and stored for subsequent analysis or immediately amplified. An aliquot of cells (5 μ l) was used as a template for the PCR reaction in a 25 μ l volume containing: 0.5 μ l KlenTaq (Clontech, Palo Alto, CA); 4.0 μ l 10 mM dNTP's (Perkin Elmer-Cetus); 2.5 μ l Kentaq buffer (Clontech); 0.25 μ l forward oligo 1; 0.25 μ l reverse oligo 2; 12.5 μ l distilled water. The sequence of the forward oligonucleotide 1 was:

5'-TGTAACACGACGGCCAGTTAAATAGACCTGCAATTATTAATCT-3' (SEQ ID NO:3)

The sequence of reverse oligonucleotide 2 was:

5'-CAGGAAACAGCTATGACCACCTGCACACCTGCAAATCCATT-3' (SEQ ID NO:4)

PCR was then performed as follows:

- | | | |
|----|---------------|------------------|
| a. | Denature | 92°C, 5 minutes |
| b. | 3 cycles of: | |
| | Denature | 92°C, 30 seconds |
| | Anneal | 59°C, 30 seconds |
| | Extend | 72°C, 60 seconds |
| c. | 3 cycles of: | |
| | Denature | 92°C, 30 seconds |
| | Anneal | 57°C, 30 seconds |
| | Extend | 72°C, 60 seconds |
| d. | 25 cycles of: | |
| | Denature | 92°C, 30 seconds |
| | Anneal | 55°C, 30 seconds |
| | Extend | 72°C, 60 seconds |

e. Hold 4°C

The underlined regions of the oligonucleotides annealed to the ADH promoter region and the amylase region, respectively, and amplified a 307 bp region from vector pSST-AMY.0 when no insert was present. Typically, the first 18 nucleotides of the 5' end of these oligonucleotides contained annealing sites for the sequencing primers. Thus, the total product of the PCR reaction from an empty vector was 343 bp. However, signal sequence-fused cDNA resulted in considerably longer nucleotide sequences.

Following the PCR, an aliquot of the reaction (5 μ l) was examined by agarose gel electrophoresis in a 1% agarose gel using a Tris-Borate-EDTA (TBE) buffering system as described by Sambrook et al., supra. Clones resulting in a single strong PCR product larger than 400 bp were further analyzed by DNA sequencing after purification with a 96 Qiaquick PCR clean-up column (Qiagen Inc., Chatsworth, CA).

EXAMPLE 3: Isolation of cDNA Clones Using Signal Algorithm Analysis

Various polypeptide-encoding nucleic acid sequences were identified by applying a proprietary signal sequence finding algorithm developed by Genentech, Inc. (South San Francisco, CA) upon ESTs as well as clustered and assembled EST fragments from public (e.g., GenBank) and/or private (LIFESEQ®, Incyte Pharmaceuticals, Inc., Palo Alto, CA) databases. The signal sequence algorithm computes a secretion signal score based on the character of the DNA nucleotides surrounding the first and optionally the second methionine codon(s) (ATG) at the 5'-end of the sequence or sequence fragment under consideration. The nucleotides following the first ATG must code for at least 35 unambiguous amino acids without any stop codons. If the first ATG has the required amino acids, the second is not examined. If neither meets the requirement, the candidate sequence is not scored. In order to determine whether the EST sequence contains an authentic signal sequence, the DNA and corresponding amino acid sequences surrounding the ATG codon are scored using a set of seven sensors (evaluation parameters) known to be associated with secretion signals. Use of this algorithm resulted in the identification of numerous polypeptide-encoding nucleic acid sequences.

EXAMPLE 4: Isolation of cDNA clones Encoding Human PRO281

In order to obtain a cDNA clone encoding PRO281, methods described in Klein et al., Proc. Natl. Acad. Sci. USA 93:7108-7113 (1996) were employed with the following modifications. Yeast transformation was performed with limiting amounts of transforming DNA in order to reduce the number of multiple transformed yeast cells. Instead of plasmid isolation from the yeast followed by transformation of *E. coli* as described in Klein et al., supra, PCR analysis was performed on single yeast colonies. PCR primers employed were bipartite in order to amplify the insert and a small portion of the invertase gene (allowing to determine that the insert was in frame with invertase) and to add on universal sequencing primer sites.

An invertase library was transformed into yeast and positives were selected on sucrose plates. Positive clones were re-tested and PCR products were sequenced. The sequence of one clone, PRO281, was determined to contain a signal peptide coding sequence. Oligonucleotide primers and probes were designed using the nucleotide sequence of PRO281. A full length plasmid library of cDNAs from human umbilical vein

endothelium tissue was titered and approximately 100,000 cfu were plated in 192 pools of 500 cfu/pool into 96-well round bottom plates. The plates were sealed and pools were grown overnight at 37°C with shaking (200rpm). PCR was performed on the individual cultures using primers. Agarose gel electrophoresis was performed and positive wells were identified by visualization of a band of the expected size. Individual positive clones were obtained by colony lift followed by hybridization with ³²P-labeled oligonucleotide. These clones were characterized by PCR, restriction digest, and southern blot analyses.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 80-82, and a stop signal at nucleotide positions 1115-1117 (Figure 1, SEQ ID NO:1). The predicted polypeptide precursor is 345 amino acids long, has a calculated molecular weight of approximately 37,205 daltons and an estimated pI of approximately 10.15. Analysis of the full-length PRO281 sequence shown in Figure 2 (SEQ ID NO:2) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 14, multiple transmembrane domains from about amino acid position 83 to about amino acid position 105, from about amino acid position 126 to about amino acid position 146, from about amino acid position 158 to about amino acid position 177, from about amino acid position 197 to about amino acid position 216, from about amino acid position 218 to about amino acid position 238, from about amino acid position 245 to about amino acid position 265, and from about amino acid position 271 to about amino acid position 290 and an amino acid sequence block having homology to G-protein coupled receptor proteins from about amino acid 115 to about amino acid 155. Clone UNQ244 (DNA16422-1209) has been deposited with ATCC on June 2, 1998 and is assigned ATCC deposit no. 209929.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 2 (SEQ ID NO:2), evidenced significant homology between the PRO281 amino acid sequence and the following Dayhoff sequences: H64634, AF033095_1, B64815, YBHL_ECOLI, EMEQUTR_1, AF064763_3, S53708, A69253, AF035413_12 and S63281.

EXAMPLE 5: Isolation of cDNA clones Encoding Human PRO276

In order to obtain a cDNA clone encoding PRO276, methods described in Klein et al., PNAS, 93:7108-7113 (1996) were employed with the following modifications. Yeast transformation was performed with limiting amounts of transforming DNA in order to reduce the number of multiple transformed yeast cells. Instead of plasmid isolation from the yeast followed by transformation of *E. coli* as described in Klein et al., supra, PCR analysis was performed on single yeast colonies. PCR primers employed were bipartite in order to amplify the insert and a small portion of the invertase gene (allowing to determine that the insert was in frame with invertase) and to add on universal sequencing primer sites.

An invertase library was transformed into yeast and positives were selected on sucrose plates. Positive clones were re-tested and PCR products were sequenced. The sequence of one clone, PRO276, was determined to contain a signal peptide coding sequence. Oligonucleotide primers and probes were designed using the nucleotide sequence of PRO276. A full length plasmid library of cDNAs from human fetal liver cells was titered and approximately 100,000 cfu were plated in 192 pools of 500 cfu/pool into 96-well round bottom

plates. The plates were sealed and pools were grown overnight at 37 C with shaking (200rpm). PCR was performed on the individual cultures using primers. Agarose gel electrophoresis was performed and positive wells were identified by visualization of a band of the expected size. Individual positive clones were obtained by colony lift followed by hybridization with ³²P-labeled oligonucleotide. These clones were characterized by PCR, restriction digest, and southern blot analyses.

5 A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 180-182 and a stop signal at nucleotide positions 933-935 (Figure 3; SEQ ID NO:5). The predicted polypeptide precursor is 251 amino acids long has a calculated molecular weight of approximately 28,801 daltons and an estimated pI of approximately 9.58. The transmembrane domains are approximately at amino acids 98-116 and 152-172 of the sequence shown in Figure 10 4 (SEQ ID NO:6). Clone DNA16435-1208 (UNQ243) has been deposited with the ATCC and is assigned ATCC deposit no. 209930 .

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 4 (SEQ ID NO:6), revealed some sequence identity between the PRO276 amino acid sequence and the following Dayhoff sequences: CEG25D7_2, 15 ATT8O5_2, S69696, GRHR_RAT, NPCBAABCD_3, AB013149_1, P_R85942 and AP000006_5.

EXAMPLE 6: Isolation of cDNA clones Encoding Human PRO189

A clone designated herein as DNA14187 was isolated as described in Example 2 above from a human retina tissue library. The DNA14187 sequence is shown in Figure 7 (SEQ ID NO:9). Based on the 20 DNA14187 sequence shown in Figure 7 (SEQ ID NO:9), oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO189. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In order to screen several libraries for a full-length clone, DNA 25 from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

A pair of PCR primers (forward and reverse) were synthesized:

30 forward PCR primer 5'-TTGACCTATACAGAGATTCATC-3' (SEQ ID NO:10); and

reverse PCR primer 5'-CTAAGAACTTCCCTCAGGATTTT-3' (SEQ ID NO:11).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA14187 sequence which had the following nucleotide sequence:

hybridization probe

5'-ATGAAGATCAATTTCAGAAGCATGCACTTCTCCTCTTGC-3' (SEQ ID NO:12).

35 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO189 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human retina tissue (LIB94). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO189 and the derived protein sequence for PRO189.

The entire nucleotide sequence of DNA21624-1391 is shown in Figure 5 (SEQ ID NO:7). Clone DNA21624-1391 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 200-202 and ending at the stop codon at nucleotide positions 1301-1303 (Figure 5). The predicted polypeptide precursor is 367 amino acids long (Figure 6). The full-length PRO189 protein shown in Figure 6 has an estimated molecular weight of about 41,871 daltons and a pI of about 5.06. Clone DNA21624-1391 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analyzing the amino acid sequence of SEQ ID NO:8, the putative N-glycosylation sites are at about amino acids 224-227, 246-249 and 285-288. A domain for cytosolic fatty-acid binding proteins is at amino acids 78-107 of SEQ ID NO:8. The corresponding nucleotides can be routinely determined given the sequences provided herein.

Some sequence identity was found to W01A6.1 and F35D11.11, C. Elegans proteins, designated in a Dayhoff database as CEW01A6_10 and CELF35D11_11, respectively. Some sequence identity was also found to an antigen to malaria and to restin, designated in a Dayhoff database as P_R05766 and AF014012_1, respectively. Some sequence identity was also found to a microtubule binding protein and to myosin, designated in a Dayhoff database as AF041382_1 and S07537, respectively. There is also some sequence identity with 1-phosphatidylinositol-4, 5-bisphosphate, designated as PIP1_RAT.

EXAMPLE 7: Isolation of cDNA clones Encoding Human PRO190

A clone designated herein as DNA14232 was isolated as described in Example 2 above from a human fetal retina tissue library. The DNA14232 sequence is shown in Figure 10 (SEQ ID NO:15). Based on the DNA14232 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO190. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CTATACCTACTGTAGCTTCT-3' (SEQ ID NO:16); and

reverse PCR primer 5'-TCAGAGAATTCCTTCCAGGA-3' (SEQ ID NO:17).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA14232 sequence which had the following nucleotide sequence:

hybridization probe

5'-ACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGCAACA-3' (SEQ ID NO:18).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO190 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human retina tissue (LIB94). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave sequences which include the full-length DNA sequence for PRO190 [herein designated as DNA23334-1392] (SEQ ID NO:13) and the derived protein sequence for PRO190.

The entire nucleotide sequence of DNA23334-1392 is shown in Figure 8 (SEQ ID NO:13). Clone DNA23334-1392 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 193-195 and which ends at the stop codon at nucleotide positions 1465-1467 (Figure 8). The predicted polypeptide precursor is 424 amino acids long (Figure 9). The full-length PRO190 protein shown in Figure 9 has an estimated molecular weight of about 48,500 daltons and a pI of about 8.65. Clone DNA23334-1392 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analyzing the amino acid sequence of SEQ ID NO:14, the putative transmembrane domains are at about amino acids 16-36, 50-74, 147-168, 229-250, 271-293, 298-318 and 328-368 of SEQ ID NO:14. N-glycosylation sites are at about amino acids 128-131, 204-207, 218-221 and 274-377 of SEQ ID NO:14. The corresponding nucleotides can be routinely determined given the sequences provided herein.

PRO190 has sequence identity with at least the following Dayhoff sequences designated as: CEZK896_2, JC5023, GMS1_SCHPO and S44668.

EXAMPLE 8: Isolation of cDNA clones Encoding Human PRO341

A clone designated herein as DNA12920 was isolated as described in Example 2 above from a human placenta tissue library. The DNA12920 sequence is shown in Figure 13 (SEQ ID NO:21). The DNA12920

sequence was then compared to various EST databases including public EST databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify homologous EST sequences. The comparison was performed using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)]. Those comparisons resulting in a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). This consensus sequence is herein designated DNA25314. Oligonucleotide primers based upon the DNA25314 sequence were then synthesized and employed to screen a human placenta cDNA library which resulted in the identification of the DNA26288-1239 clone shown in Figure 11. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 380-382, and a stop signal at nucleotide positions 1754-1756 (Figure 11, SEQ ID NO:19). The predicted polypeptide precursor is 458 amino acids long, has a calculated molecular weight of approximately 50,264 daltons and an estimated pI of approximately 8.17. Analysis of the full-length PRO341 sequence shown in Figure 12 (SEQ ID NO:20) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 17, transmembrane domains from about amino acid 171 to about amino acid 190, from about amino acid 220 to about amino acid 239, from about amino acid 259 to about amino acid 275, from about amino acid 286 to about amino acid 305, from about amino acid 316 to about amino acid 335, from about amino acid 353 to about amino acid 378 and from about amino acid 396 to about amino acid 417 and potential N-glycosylation sites from about amino acid 145 to about amino acid 147 and from about amino acid 155 to about amino acid 158. Clone DNA26288-1239 has been deposited with ATCC on April 21, 1998 and is assigned ATCC deposit no. 209792.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 12 (SEQ ID NO:20), evidenced homology between the PRO341 amino acid sequence and the following Dayhoff sequences: S75696, H69788, D69852, A69888, B64918, F64752, LPU89276_1, G64962, S52977 and S44253.

EXAMPLE 9: Isolation of cDNA clones Encoding Human PRO180

A clone designated herein as DNA12922 was isolated as described in Example 2 above from a human placenta tissue library. The DNA12922 sequence is shown in Figure 16 (SEQ ID NO:24). The DNA12922 sequence was then compared to various EST databases including public EST databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify homologous EST sequences. The comparison was performed using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)]. Those comparisons resulting in a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington).

An oligonucleotide probe was formed based upon the consensus sequence obtained above. This probe had the following sequence.

5'-ACCTGTTAGAAATGTGGTGGTTTCAGCAAGGCCTCAGTTT (SEQ ID NO:25).

This probe was used to screen a human placenta library prepared as described in paragraph 1 of Example 2 above. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, 253:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp. A clone designated herein as DNA26843-1389 was obtained.

The entire nucleotide sequence of DNA26843-1389 is shown in Figure 14 (SEQ ID NO:22). Clone DNA26843-1389 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 121-123 and ending at the stop codon at nucleotide positions 919-921 (Figure 14). The predicted polypeptide precursor is 266 amino acids long (Figure 15). The full-length PRO180 protein shown in Figure 15 has an estimated molecular weight of about 29,766 daltons and a pI of about 8.39. Clone DNA26843-1389 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Still analyzing the amino acid sequence of SEQ ID NO:23, the transmembrane domains are at about amino acids 13-33 (type II), 54-73, 94-113, 160-180 and 122-141 of SEQ ID NO:23. N-myristoylation sites are at about amino acids 57-62, 95-100, 99-104, 124-129 and 183-188 of SEQ ID NO:23. The corresponding nucleotides can be routinely determined given the sequences provided herein.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 15 (SEQ ID NO:23), evidenced some sequence identity between the PRO180 amino acid sequence and the following Dayhoff sequences: CEC33A11_2, CEG11E6_5, CELW03A5_1 AND PEU83861_2 (NADH dehydrogenase subunit 4L, mitochondrion).

EXAMPLE 10: Isolation of cDNA clones Encoding Human PRO194

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein DNA19464. Based on the DNA19464 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO194. PCR primers (forward and reverse) were synthesized based upon the DNA19464 sequence. Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA19464 sequence.

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO194 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB25).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO194 [herein designated as DNA26844-1394] (SEQ ID NO:27) and the derived protein sequence for PRO194.

The entire nucleotide sequence of DNA26844-1394 is shown in Figure 17 (SEQ ID NO:27). Clone DNA26844-1394 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 81-83 and ending at the stop codon at nucleotide positions 873-875 (Figure 17). The predicted polypeptide precursor is 264 amino acids long (Figure 18). The full-length PRO194 protein shown in Figure 18 has an estimated molecular weight of about 29,665 daltons and a pI of about 9.34. Analysis of the full-length PRO194 sequence shown in Figure 18 (SEQ ID NO:28) evidences the presence of various important polypeptides domains as shown in Figure 18. Clone DNA26844-1394 has been deposited with ATCC on June 2, 1998 and is assigned ATCC deposit no. 209926.

Analysis of the amino acid sequence of the full-length PRO194 polypeptide suggests that it does not exhibit significant sequence similarity to any known human protein. However, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some homology between the PRO194 amino acid sequence and the following Dayhoff sequences, HUMORFT_1, CET07F10_5, ATFCA9_12, F64934, YDJX_ECOLI, ATAF00065719F29G20.19, H70002, S76980, H64934 and S76385.

EXAMPLE 11: Isolation of cDNA clones Encoding Human PRO203

A clone designated herein as DNA15618 was isolated as described in Example 2 above from a human fetal lung tissue library. The DNA15618 sequence is shown in Figure 21 (SEQ ID NO:31). Oligonucleotide probes were generated from the sequence of the DNA15618 molecule and were used to screen a human fetal lung library (LIB26) prepared as described in paragraph 1 of Example 2 above. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 159-161 and ending at the stop codon found at nucleotide positions 1200-1202 (Figure 19; SEQ ID NO:29). The predicted polypeptide precursor is 347 amino acids long, has a calculated molecular weight of approximately 39,870 daltons and an estimated pI of approximately 6.76. Analysis of the full-length PRO203 sequence shown in Figure 20 (SEQ ID NO:30) evidences the presence of the following: a type II transmembrane domain at about amino acid 64 to about amino acid 87; possible N-glycosylation sites at about amino acid 147 to about amino acid 150, about amino acid 155 to about amino acid 158, and about amino acid 237 to about amino acid 240; sequence identity with heavy-metal-associated domain proteins at about amino acid 23 to about amino acid 45, and sequence identity with D-isomer specific 2-hydroxyacid dehydrogenase at about amino acid 24 to about amino acid 34. Clone DNA30862-1396 was deposited with the ATCC on June 2, 1998, and is assigned ATCC deposit no. 209920.

Analysis of the amino acid sequence of the full-length PRO203 polypeptide suggests that it possesses sequence similarity to GST ATPase, thereby indicating that PRO203 may be a novel GST ATPase. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO203 amino acid sequence and the following Dayhoff sequences, AF008124_1, CFRCD1GEN_1, and P_R82566.

EXAMPLE 12: Isolation of cDNA clones Encoding Human PRO290

An expressed sequence tag (EST) DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST was identified that had homology to beige and FAN. An oligonucleotide probe based upon the identified EST sequence was then synthesized and used to screen human fetal kidney cDNA libraries in an attempt to identify a full-length cDNA clone. The oligonucleotide probe had the following sequence:

5' TGACTGCACTACCCCGTGGCAAGCTGTTGAGCCAGCTCAGCTG 3' (SEQ ID NO:34).

RNA for construction of cDNA libraries was isolated from human fetal kidney tissue. The cDNA libraries used to isolate the cDNA clones encoding human PRO290 were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science* 253:1278-1280 (1991)) in the unique XhoI and NotI.

A cDNA clone was identified and sequenced in entirety. The entire nucleotide sequence of DNA35680-1212 is shown in Figure 22 (SEQ ID NO:32). Clone DNA35680-1212 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 293-295, and a stop codon at nucleotide positions 3302-3304 (Figure 22; SEQ ID NO:32). The predicted polypeptide precursor is 1003 amino acids long.

It is currently believed that the PRO290 polypeptide is related to FAN and/or beige. Clone DNA35680-1212 has been deposited with ATCC and is assigned ATCC deposit no. 209790. It is understood that the deposited clone has the actual correct sequence rather than the representations provided herein. The full-length PRO290 protein shown in Figure 23 has an estimated molecular weight of about 112,013 daltons and a pI of about 6.4.

EXAMPLE 13: Isolation of cDNA Clones Encoding Human PRO874

A consensus DNA sequence designated herein as DNA36459 was identified using phrap as described in Example 1 above. Based on the DNA36459 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the coding sequence for PRO874.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TCGTGCCCAGGGGCTGATGTGC-3' (SEQ ID NO:37); and

reverse PCR primer 5'-GTCTTTACCCAGCCCCGGGATGCG-3' (SEQ ID NO:38).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA36459 sequence which had the following nucleotide sequence:

hybridization probe

5'-GGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCTGGG-3' (SEQ ID NO:39).

In order to screen several libraries for a source of a clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate

clones encoding the PRO874 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB25).

DNA sequencing of the clones isolated as described above gave the DNA sequence for PRO874 [herein designated as DNA40621-1440] (SEQ ID NO:35) and the derived protein sequence for PRO874.

The entire nucleotide sequence of DNA40621-1440 is shown in Figure 24 (SEQ ID NO:35). Clone DNA40621-1440 contains a single open reading frame ending at the stop codon at nucleotide positions 964-966 (Figure 24). The predicted polypeptide encoded by DNA40621-1440 is 321 amino acids long (Figure 25). The PRO874 protein shown in Figure 25 has an estimated molecular weight of about 36,194 daltons and a pI of about 9.85. Analysis of the PRO874 sequence shown in Figure 25 (SEQ ID NO:36) evidenced the presence of the following: a type II transmembrane domain at about amino acids 57-80; additional transmembrane domains at about amino acids 110-126, 215-231, and 254-274; potential N-glycosylation sites at about amino acids 16-19, 27-30, and 289-292; sequence identity with hypothetical YBR002c family proteins at about amino acids 276-287; and sequence identity with ammonium transporter proteins at about amino acids 204-230. Clone DNA40621-1440 was deposited with the ATCC on June 2, 1998, and is assigned ATCC deposit no. 209922.

Analysis of the amino acid sequence of the PRO874 polypeptide suggests that it is a novel multi-span transmembrane protein. However, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced sequence identity between the PRO874 amino acid sequence and the following Dayhoff sequences: S67049, AF054839_1, S73437, S52460, and HIVU80570_1.

EXAMPLE 14: Isolation of cDNA Clones Encoding Human PRO710

A yeast screening assay was employed to identify cDNA clones that encoded potential secreted proteins. Use of this yeast screening assay allowed identification of a single cDNA clone whose sequence (herein designated as DNA38190) is shown in Figure 28 (SEQ ID NO:42). Based on the DNA38190 sequence shown in Figure 28, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO710. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TTCCGCAAAGAGTTCTACGAGGTGG-3' (SEQ ID NO:43)

reverse PCR primer 5'-ATTGACAACATTGACTGGCCTATGGG-3' (SEQ ID NO:44)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA38190 sequence which had the following nucleotide sequence

hybridization probe

5'-GTGGATGCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTCCAGTGTGA-3' (SEQ ID NO:45)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO710 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 67-69 and ending at the stop codon found at nucleotide positions 1765-1767 (Figure 26, SEQ ID NO:40). The predicted polypeptide precursor is 566 amino acids long, has a calculated molecular weight of approximately 65,555 daltons and an estimated pI of approximately 5.44. Analysis of the full-length PRO710 sequence shown in Figure 27 (SEQ ID NO:41) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 32, a transmembrane domain from about amino acid 454 to about amino acid 476, an aminoacyl-transfer RNA synthetase class-II signature sequence from about amino acid 6 to about amino acid 26 and potential N-glycosylation sites from about amino acid 111 to about amino acid 114, from about amino acid 146 to about amino acid 149 and from about amino acid 292 to about amino acid 295. Clone DNA44161-1434 has been deposited with ATCC on May 27, 1998 and is assigned ATCC deposit no. 209907.

Analysis of the amino acid sequence of the full-length PRO710 polypeptide suggests that it possesses significant sequence similarity to the CDC45 protein, thereby indicating that PRO710 may be a novel CDC45 homolog. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO710 amino acid sequence and the following Dayhoff sequences, HSAJ3728_1, CEF34D10_1, S64939, UMU50276_1, TRHY_SHEEP, CELT14E8_1, RNA1_YEAST, LVU89340_1, HSU80736_1 and CEZK337_2.

EXAMPLE 15: Isolation of cDNA clones Encoding Human PRO1151

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA40665. Based on the DNA40665 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1151.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCAGACGCTGCTCTTCGAAAGGGTC-3' (SEQ ID NO:48)

reverse PCR primer 5'-GGTCCCCGTAGGCCAGGTCCAGC-3' (SEQ ID NO:49)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA40665 sequence which had the following nucleotide sequence

hybridization probe

5'-CTACTTCTTCAGCCTCAATGTGCACAGCTGGAATTACAAGGAGACGTACG-3' (SEQ ID NO:50)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1151 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1151 (designated herein as DNA44694-1500 [Figure 29, SEQ ID NO:46]; and the derived protein sequence for PRO1151.

The entire nucleotide sequence of DNA44694-1500 is shown in Figure 29 (SEQ ID NO:46). Clone DNA44694-1500 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 272-274 and ending at the stop codon at nucleotide positions 1049-1051 (Figure 29). The predicted polypeptide precursor is 259 amino acids long (Figure 30). The full-length PRO1151 protein shown in Figure 30 has an estimated molecular weight of about 28,770 daltons and a pI of about 6.12. Analysis of the full-length PRO1151 sequence shown in Figure 30 (SEQ ID NO:47) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 20, a potential N-glycosylation site from about amino acid 72 to about amino acid 75 and amino acid sequence blocks having homology to C1q domain-containing proteins from about amino acid 144 to about amino acid 178, from about amino acid 78 to about amino acid 111 and from about amino acid 84 to about amino acid 117. Clone UNQ581 (DNA44694-1500) has been deposited with ATCC on August 11, 1998 and is assigned ATCC deposit no. 203114.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 30 (SEQ ID NO:47), evidenced significant homology between the PRO1151 amino acid sequence and the following Dayhoff sequences: ACR3_HUMAN, HP25_TAMAS, HUMC1QB2_1, P_R99306, CA1F_HUMAN, JX0369, CA24_HUMAN, S32436, P_R28916 and CA54_HUMAN.

EXAMPLE 16: Isolation of cDNA clones Encoding Human PRO1282

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein as DNA33778. Based on the DNA33778 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1282.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'TCTTCAGCCGCTTGCGCAACCTC3' (SEQ ID NO:53); and

reverse PCR primer 5'TTGCTCACATCCAGCTCCTGCAGG3' (SEQ ID NO:54).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA33778 sequence which had the following nucleotide sequence:

hybridization probe

5'TGGATGTTGTCCAGACAACCAGCTGGAGCTGTATCCGAGGC3' (SEQ ID NO:55).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1282 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal liver.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1282 (designated herein as DNA45495-1550 [Figure 31, SEQ ID NO:51]; and the derived protein sequence for PRO1282.

The entire coding sequence of PRO1282 is shown in Figure 31 (SEQ ID NO:51). Clone DNA45495-1550 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 120-122, and an apparent stop codon at nucleotide positions 2139-2141 (SEQ ID NO:51). The predicted polypeptide precursor is 673 amino acids long. The signal peptide is at about amino acids 1-23; the transmembrane domain is at about amino acids 579-599; an EGF-like domain cysteine pattern signature starts at about amino acid 430; and leucine zipper patterns start at about amino acids 197 and 269 of SEQ ID NO:52, see Figure 32. Clone DNA45495-1550 has been deposited with the ATCC and is assigned ATCC deposit no. 203156. The full-length PRO1282 protein shown in Figure 32 has an estimated molecular weight of about 71,655 daltons and a pI of about 7.8.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 32 (SEQ ID NO:52), revealed sequence identity between the PRO1282 amino acid sequence and the following Dayhoff sequences (data from database incorporated by reference): AB007876_1, RNPLGPV_1, MUSLRRP_1, ALS_PAPPA, AC004142_1, ALS_HUMAN, AB014462_1, DMTARTAN_1, HSCHON03_1 and S46224.

EXAMPLE 17: Isolation of cDNA clones Encoding Human PRO358

Using the method described in Example 1 above, a single EST sequence was identified in the Incyte database, designated herein as INC3115949. Based on the INC3115949 EST sequence, oligonucleotides were synthesized to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for PRO358.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TCCCACCAGGTATCATAACTGAA-3' (SEQ ID NO:58)

reverse PCR primer 5'-TTATAGACAATCTGTTCTCATCAGAGA-3' (SEQ ID NO:59)

A probe was also synthesized:

5'-AAAAAGCATACTTGAATGGCCCAAGGATAGGTGTAAATG-3' (SEQ ID NO:60)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used

to isolate clones encoding the PRO358 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human bone marrow (LIB256). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO358 (Figure 33, SEQ ID NO:56) and the derived protein sequence for PRO358 (Figures 34, SEQ ID NO:57).

The entire nucleotide sequence of the clone identified (DNA47361-1154) is shown in Figure 33 (SEQ ID NO:56). Clone DNA47361-1154 contains a single open reading frame with an apparent translational initiation site (ATG start signal) at nucleotide positions underlined in Figure 33. The predicted polypeptide precursor is 811 amino acids long, including a putative signal sequence (amino acids 1 to 19), an extracellular domain (amino acids 20 to 575, including leucine rich repeats in the region from position 55 to position 575), a putative transmembrane domain (amino acids 576 to 595). Clone DNA47361-1249 has been deposited with ATCC and is assigned ATCC deposit no. 209431.

EXAMPLE 18: Isolation of cDNA clones Encoding Human PRO1310

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein as DNA37164. Based on the DNA37164 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1310.

PCR primers (forward and reverse) were synthesized:
forward PCR primer: 5'GTTCTCAATGAGCTACCCGTCCCC3' (SEQ ID NO:63) and
reverse PCR primer: 5'CGCGATGTAGTGGAAGCTCGGGCTC3' (SEQ ID NO:64).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA47394 sequence which had the following nucleotide sequence:

hybridization probe:
5'ATCCGCATAAACCCCTCAGTCCTGGTTTGATAATGGGAGCATCTGCATGAG3' (SEQ ID NO:65).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1310 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal liver tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1310 and the derived protein sequence for PRO1310.

The entire coding sequence of PRO1310 is shown in Figure 35 (SEQ ID NO:61). Clone DNA47394-1572 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 326-328, and an apparent stop codon at nucleotide positions 2594-2596 (SEQ ID NO:61). The predicted polypeptide precursor is 765 amino acids long. The signal peptide is at about amino acids 1-25 of SEQ ID NO:62. Clone DNA47394-1572 has been deposited with ATCC and is assigned ATCC deposit no. 203109.

5 The full-length PRO1310 protein shown in Figure 36 has an estimated molecular weight of about 85,898 daltons and a pI of about 6.87.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 36 (SEQ ID NO:62), revealed sequence identity between the PRO1310 amino acid sequence and the following Dayhoff sequences: AF017639_1, P_W36817, JC5256, CBPH_HUMAN, MMU23184_1, CBPN_HUMAN, HSU83411_1, CEF01D4_7, RNU62897_1 and P_W11851.

10

EXAMPLE 19: Isolation of cDNA Clones Encoding Human PRO698

A yeast screening assay was employed to identify cDNA clones that encoded potential secreted proteins. Use of this yeast screening assay allowed identification of a single cDNA clone whose sequence (herein designated as DNA39906) is shown in Figure 39 (SEQ ID NO:68). Based on the DNA39906 sequence shown in Figure 39, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO698. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by

15

20 PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AGCTGTGGTCATGGTGGTGTGGTG-3' (SEQ ID NO:69)

reverse PCR primer 5'-CTACCTTGCCATAGGTGATCCGC-3' (SEQ ID NO:70)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA39906 sequence which had the following nucleotide sequence

hybridization probe

5'-CATCAGCAAACCGTCTGTGGTTCAGCTCAACTGGAGAGGGTT-3' (SEQ ID NO:71)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO698 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human bone marrow tissue (LIB255). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or

30

35

pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 14-16 and ending at the stop codon found at nucleotide positions 1544-1546 (Figure 37, SEQ ID NO:66). The predicted polypeptide precursor is 510 amino acids long, has a calculated molecular weight of approximately 57,280 daltons and an estimated pI of approximately 5.61. Analysis of the full-length PRO698 sequence shown in Figure 38 (SEQ ID NO:67) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 20, potential N-glycosylation sites from about amino acid 72 to about amino acid 75, from about amino acid 136 to about amino acid 139, from about amino acid 193 to about amino acid 196, from about amino acid 253 to about amino acid 256, from about amino acid 352 to about amino acid 355 and from about amino acid 411 to about amino acid 414 an amino acid block having homology to legume lectin beta-chain proteins from about amino acid 20 to about amino acid 39 and an amino acid block having homology to the HBGF/FGF family of proteins from about amino acid 338 to about amino acid 365. Clone DNA48320-1433 has been deposited with ATCC on May 27, 1998 and is assigned ATCC deposit no. 209904.

Analysis of the amino acid sequence of the full-length PRO698 polypeptide suggests that it possesses significant sequence similarity to the olfactomedin protein, thereby indicating that PRO698 may be a novel olfactomedin homolog. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO698 amino acid sequence and the following Dayhoff sequences, OLFM_RANCA, I73637, AB006686S3_1, RNU78105_1, RNU72487_1, P_R98225, CELC48E7_4, CEF11C3_3, XLU85970_1 and S42257.

EXAMPLE 20: Isolation of cDNA Clones Encoding Human PRO732

A yeast screening assay was employed to identify cDNA clones that encoded potential secreted proteins. Use of this yeast screening assay allowed identification of a single cDNA clone whose sequence (herein designated as DNA42580) is shown in Figure 45 (SEQ ID NO:77). The DNA42580 sequence was then compared to a variety of known EST sequences to identify homologies. The EST databases employed included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison to a 6 frame translation of the EST sequence. Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington).

Using the above analysis, a consensus DNA sequence was assembled relative to other EST sequences using phrap. This consensus sequence is herein designated consen01. Proprietary Genentech EST sequences were employed in the consensus assembly and they are herein designated DNA20239 (Figure 42; SEQ ID NO:74), DNA38050 (Figure 43; SEQ ID NO:75) and DNA40683 (Figure 44; SEQ ID NO:76).

Based on the consen01 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO732. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-ATGTTTGTGTGGAAGTGCCCCG-3' (SEQ ID NO:78)

forward PCR primer 5'-GTCAACATGCTCCTCTGC-3' (SEQ ID NO:79)

reverse PCR primer 5'-AATCCATTGTGCACTGCAGCTCTAGG-3' (SEQ ID NO:80)

reverse PCR primer 5'-GAGCATGCCACCACTGGACTGAC-3' (SEQ ID NO:81)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA44143 sequence which had the following nucleotide sequence

hybridization probe

5'-GCCGATGCTGTCCTAGTGGAACAACCTCCACTGTAAGTATGATCTATGCAC-3' (SEQ ID NO:82)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO732 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB26). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 88-90 and ending at the stop codon found at nucleotide positions 1447-1449 (Figure 40, SEQ ID NO:72). The predicted polypeptide precursor is 453 amino acids long, has a calculated molecular weight of approximately 50,419 daltons and an estimated pI of approximately 5.78. Analysis of the full-length PRO732 sequence shown in Figure 41 (SEQ ID NO:73) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 28, transmembrane domains from about amino acid 37 to about amino acid 57, from about amino acid 93 to about amino acid 109, from about amino acid 126 to about amino acid 148, from about amino acid 151 to about amino acid 172, from about amino acid 197 to about amino acid 215, from about amino acid 231 to about amino acid 245, from about

amino acid 260 to about amino acid 279, from about amino acid 315 to about amino acid 333, from about amino acid 384 to about amino acid 403 and from about amino acid 422 to about amino acid 447, potential N-glycosylation sites from about amino acid 33 to about amino acid 36, from about amino acid 34 to about amino acid 37, from about amino acid 179 to about amino acid 183, from about amino acid 298 to about amino acid 301, from about amino acid 337 to about amino acid 340 and from about amino acid 406 to about amino acid 409, an amino acid block having homology to the MIP family of proteins from about amino acid 119 to about amino acid 149 and an amino acid block having homology to DNA/RNA non-specific endonuclease proteins from about amino acid 279 to about amino acid 286. Clone DNA48334-1435 has been deposited with ATCC on June 2, 1998 and is assigned ATCC deposit no. 209924.

Analysis of the amino acid sequence of the full-length PRO732 polypeptide suggests that it possesses significant sequence similarity to the Diff33 protein, thereby indicating that PRO732 may be a novel Diff33 homolog. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO732 amino acid sequence and the following Dayhoff sequences, HS179M20_2, MUSTETU_1, CER11H6_2, RATDRP_1, S51256, E69226, AE000869_1, JC4120, CYB_PARTE and P_R50619.

EXAMPLE 21: Isolation of cDNA clones Encoding Human PRO1120

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein consen0352. The consen0352 sequence was then extended using repeated cycles of BLAST and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above. The extended consensus sequence is designated herein as DNA34365. Based on the DNA34365 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1120.

PCR primers (forward and reverse) were synthesized:

forward PCR primers: 5'-GAAGCCGGCTGTCTGAATC-3' (SEQ ID NO:85),
5'-GGCCAGCTATCTCCGCAG-3' (SEQ ID NO:86), 5'-AAGGGCCTGCAAGAGAAG-3' (SEQ ID NO:87), 5'-CACTGGGACAACGTGTGGG-3' (SEQ ID NO:88),
5'-CAGAGGCAACGTGGAGAG-3' (SEQ ID NO:89), and
5'-AAGTATTGTCATACAGTGTTTC-3' (SEQ ID NO:90);
reverse PCR primers: 5'-TAGTACTTGGGCACGAGGTTGGAG-3' (SEQ ID NO:91), and 5'-TCATACCAACTGCTGGTCATTGGC-3' (SEQ ID NO:92).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA34365 consensus sequence which had the following nucleotide sequence:

hybridization probe:

5'-CTCAAGCTGCTGGACACGGAGCGGCCGGTGAATCGGTTTCACTTG-3' (SEQ ID NO:93).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used

to isolate clones encoding the PRO1120 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1120 (designated herein as DNA48606-1479 [Figure 46, SEQ ID NO:83]; and the derived protein sequence for PRO1120.

The entire coding sequence of PRO1120 is shown in Figure 46 (SEQ ID NO:83). Clone DNA48606-1479 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 608-610 and an apparent stop codon at nucleotide positions 3209-3211. The predicted polypeptide precursor is 867 amino acids long. The full-length PRO1120 protein shown in Figure 47 has an estimated molecular weight of about 100,156 Daltons and a pI of about 9.44. Additional features of the PRO1120 polypeptide include a signal peptide at about amino acids 1-17; a sulfatase signature at about amino acids 86-98; regions of homology to sulfatases at about amino acids 87-106, 133-146, 216-229, 291-320, and 365-375; and potential N-glycosylation sites at about amino acids 65-68, 112-115, 132-135, 149-152, 171-174, 198-201, 241-245, 561-564, 608-611, 717-720, 754-757, and 764-767.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 47 (SEQ ID NO:84), revealed significant homology between the PRO1120 amino acid sequence and the following Dayhoff sequences: CELK09C4_1, GL6S_HUMAN, G65169, NCU89492_1, BCU44852_1, E64903, P_R51355, STS_HUMAN, GA6S_HUMAN, and IDS_MOUSE. Clone DNA48606-1479 was deposited with the ATCC on July 1, 1998, and is assigned ATCC deposit no. 203040.

EXAMPLE 22: Isolation of cDNA clones Encoding Human PRO537

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated as Incyte EST cluster no. 29605. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA48350.

In light of an observed sequence homology between the DNA48350 consensus sequence and an EST sequence encompassed within the Merck EST clone no. R63443, the Merck EST clone R63443 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 48 and is herein designated as DNA49141-1431.

Clone DNA49141-1431 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 97-99 and ending at the stop codon at nucleotide positions 442-444 (Figure 48).

The predicted polypeptide precursor is 115 amino acids long (Figure 49). The full-length PRO537 protein shown in Figure 49 has an estimated molecular weight of about 13,183 daltons and a pI of about 12.13. Analysis of the full-length PRO537 sequence shown in Figure 49 (SEQ ID NO:95) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 31, a potential N-glycosylation site from about amino acid 44 to about amino acid 47, potential N-myristolation sites from about amino acid 3 to about amino acid 8 and from about amino acid 16 to about amino acid 21 and an amino acid block having homology to multicopper oxidase proteins from about amino acid 97 to about amino acid 105. Clone DNA49141-1431 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203003.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 49 (SEQ ID NO:95), evidenced homology between the PRO537 amino acid sequence and the following Dayhoff sequences: A54523, CELF22H10_2, FKH4_MOUSE, OTX1_HUMAN, URB1_USTMA, KNOB_PLAFN, A32895_1, AF036332_1, HRG_HUMAN and HRP3_PLAFS.

EXAMPLE 23: Isolation of cDNA clones Encoding Human PRO536

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated herein as ss.clu2437.init. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA48351.

In light of an observed sequence homology between the DNA48351 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H11129, the Merck EST clone H11129 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 50 and is herein designated as DNA49142-1430.

Clone DNA49142-1430 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 48-50 and ending at the stop codon at nucleotide positions 987-989 (Figure 50). The predicted polypeptide precursor is 313 amino acids long (Figure 51). The full-length PRO536 protein shown in Figure 51 has an estimated molecular weight of about 34,189 daltons and a pI of about 4.8. Analysis of the full-length PRO536 sequence shown in Figure 51 (SEQ ID NO:97) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 25, a potential N-glycosylation site from about amino acid 45 to about amino acid 48 and an amino acid sequence block having homology to sulfatase proteins from about amino acid 16 to about amino acid 26. Clone DNA49142-1430 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203002.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 51 (SEQ ID NO:97), evidenced homology between the PRO536 amino acid sequence and the following Dayhoff sequences: APU46857_1, PK2_DICDI, H64743, F5I14_18, CEAM_ECOLI, GEN14267, H64965, TCU39815_1, PSBJ_ODOSI and P_R06980.

5 EXAMPLE 24: Isolation of cDNA clones Encoding Human PRO535

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated herein as ss.clu12694.init. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA48352. Two proprietary Genentech EST sequences were employed in the assembly are herein shown in Figures 54 and 55.

In light of an observed sequence homology between the DNA48352 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H86994, the Merck EST clone H86994 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 52 and is herein designated as DNA49143-1429.

Clone DNA49143-1429 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 78-80 and ending at the stop codon at nucleotide positions 681-683 (Figure 52). The predicted polypeptide precursor is 201 amino acids long (Figure 53). The full-length PRO535 protein shown in Figure 53 has an estimated molecular weight of about 22,180 daltons and a pI of about 9.68. Analysis of the full-length PRO535 sequence shown in Figure 53 (SEQ ID NO:99) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 25, a transmembrane domain from about amino acid 155 to about amino acid 174, a potential N-glycosylation site from about amino acid 196 to about amino acid 199 and FKBP-type peptidyl-prolyl cis-trans isomer signature sequences from about amino acid 62 to about amino acid 77, from about amino acid 87 to about amino acid 123 and from about amino acid 128 to about amino acid 141. Clone DNA49143-1429 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203013.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST- sequence alignment analysis of the full-length sequence shown in Figure 53 (SEQ ID NO:99), evidenced homology between the PRO535 amino acid sequence and the following Dayhoff sequences: S71237, P_R93551, P_R28980, S71238, FKB2_HUMAN, CELC05C8_1, S55383, S72485, CELC50F2_6 and S75144.

EXAMPLE 25: Isolation of cDNA clones Encoding Human PRO718

A cDNA sequence isolated in the amylase screen described in Example 2 (human fetal lung library) above is herein designated DNA43512 (see Figure 62; SEQ ID NO:108). The DNA43512 sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA45625. Proprietary Genentech EST sequences were employed in the assembly and are herein shown in Figures 58-61.

Based on the DNA45625 sequence, oligonucleotide probes were generated and used to screen a human fetal lung library (LIB25) prepared as described in paragraph 1 of Example 2 above. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GGGTGGATGGTACTGCTGCATCC-3' (SEQ ID NO:109)

reverse PCR primer 5'-TGTTGTGCTGTGGGAAATCAGATGTG-3' (SEQ ID NO:110)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA45625 sequence which had the following nucleotide sequence:

hybridization probe

5'-GTGTCTGGAGGCTGTGGCCGTTTGTGTTTCTTGGGCTAAAATCGGG-3' (SEQ ID NO:111)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO718 gene using the probe oligonucleotide and one of the PCR primers.

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 36-38 and ending at the stop codon found at nucleotide positions 607-609 (Figure 56; SEQ ID NO:102). The predicted polypeptide precursor is 157 amino acids long, has a calculated molecular weight of approximately 17,400 daltons and an estimated pI of approximately 5.78. Analysis of the full-length PRO718 sequence shown in Figure 57 (SEQ ID NO:103) evidences the presence of the following: a type II transmembrane domain from about amino acid 21 to about amino acid 40, and other transmembrane domains at about amino acid 58 to about amino acid 78, about amino acid 95 to about amino acid 114, and about amino acid 127 to about amino acid 147; a cell attachment sequence from about amino acid 79 to about amino acid 81; and a potential N-glycosylation site from about amino acid 53 to about amino acid 56. Clone DNA49647-1398 has been deposited with ATCC on June 2, 1998 and is assigned ATCC deposit no. 209919.

Analysis of the amino acid sequence of the full-length PRO718 polypeptide suggests that it possesses no significant sequence similarity to any known protein. However, an analysis of the Dayhoff database

(version 35.45 SwissProt 35) evidenced some degree of homology between the PRO718 amino acid sequence and the following Dayhoff sequences: AF045606_1, AF039906_1, SPBC8D2_2, S63441, F64728, COX1_TRYBB, F64375, E64173, RPYGJT_3, MTCY261_23.

EXAMPLE 26: Isolation of cDNA clones Encoding Human PRO872

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST sequence designated herein as clu120709.init. The clu120709.init sequence was then compared a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA48254.

In light of an observed sequence homology between the DNA48254 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3438068, the Incyte EST clone 3438068 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 63 and is the full-length DNA sequence for PRO872. Clone DNA49819-1439 was deposited with the ATCC on June 2, 1998, and is assigned ATCC deposit no. 209931.

The entire nucleotide sequence of DNA49819-1439 is shown in Figure 63 (SEQ ID NO:112). Clone DNA49819-1439 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 14-16 and ending at the stop codon at nucleotide positions 1844-1846 (Figure 63). The predicted polypeptide precursor is 610 amino acids long (Figure 64). The full-length PRO872 protein shown in Figure 64 has an estimated molecular weight of about 66,820 daltons and a pI of about 8.65. Analysis of the full-length PRO872 sequence shown in Figure 64 (SEQ ID NO:113) evidences the presence of the following features: a signal peptide at amino acid 1 to about 18, putative transmembrane domains at about amino acids 70-87, 200-222 and 568-588; sequence identity with bacterial-type phytoene dehydrogenase protein at about amino acids 71-105; sequence identity with a regulator of chromosome condensation (RCC1) signature 2 at about amino acids 201-211; leucine zipper patterns at about amino acids 214-235, 221-242, 228-249 and 364-385; a potential N-glycosylation site at about amino acids 271-274; and a glycosaminoglycan attachment site at about amino acids 75-78. Analysis of the amino acid sequence of the full-length PRO872 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO872 amino acid sequence and the following Dayhoff sequences: PRCRTI_1, S75951, S74689, CELF37C4_3, CRTI_RHOCA, S76617, YNI2_METTL, MTV014_14, AOFB_HUMAN, and MMU70429_1.

EXAMPLE 27: Isolation of cDNA clones Encoding Human PRO1063

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence designated herein as ss.clu119743.init. The Incyte EST cluster sequence

ss.clu119743.init sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA48288.

In light of an observed sequence homology between the DNA48288 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2783726, the Incyte EST clone 2783726 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 65 and is herein designated DNA49820-1427.

The full length clone shown in Figure 65 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 90-92 and ending at the stop codon found at nucleotide positions 993-995 (Figure 65; SEQ ID NO:114). The predicted polypeptide precursor is 301 amino acids long, has a calculated molecular weight of approximately 33,530 daltons and an estimated pI of approximately 4.80. Analysis of the full-length PRO1063 sequence shown in Figure 66 (SEQ ID NO:115) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 21, potential N-glycosylation sites from about amino acid 195 to about amino acid 198, from about amino acid 217 to about amino acid 220 and from about amino acid 272 to about amino acid 275, a glycosaminoglycan attachment site from about amino acid 267 to about amino acid 270, a microbodies C-terminal targeting signal site from about amino acid 299 to about amino acid 301, a type II fibronectin collagen-binding domain homology sequence from about amino acid 127 to about amino acid 168 and a fructose-bisphosphate aldolase class II protein homology sequence from about amino acid 101 to about amino acid 118. Clone DNA49820-1427 has been deposited with the ATCC on June 2, 1998 and is assigned ATCC deposit no. 209932.

Analysis of the amino acid sequence of the full-length PRO1063 polypeptide suggests that it possesses sequence similarity to the human type IV collagenase protein. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some degree of homology between the PRO1063 amino acid sequence and the following Dayhoff sequences, S68303, CFU68533_1, P_P91139, RNU65656_1, PA2R_RABIT, MMU56734_1, FINC_XENLA, A48925, P_R92778 and FA12_HUMAN.

EXAMPLE 28: Isolation of cDNA clones Encoding Human PRO619

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated herein as 88434. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a

BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington).

In light of an observed sequence homology between the consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 1656694, the Incyte EST clone 1656694 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 67 and is herein designated as DNA49821-1562.

The full length clone shown in Figure 67 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 81-83 and ending at the stop codon found at nucleotide positions 450-452 (Figure 67; SEQ ID NO:116). The predicted polypeptide precursor (Figure 68, SEQ ID NO:117) is 123 amino acids long including a predicted signal peptide at about amino acids 1-20. PRO619 has a calculated molecular weight of approximately 13,710 daltons and an estimated pI of approximately 5.19. Clone DNA49821-1562 was deposited with the ATCC on June 16, 1998 and is assigned ATCC deposit no. 209981.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 68 (SEQ ID NO:117), revealed significant homology between the PRO619 amino acid sequence and the following Dayhoff sequences: S35302, D87009_1, HSU93494_1, HUMIGLAM5_1, D86999_2, HUMIGLYM1_1, HUMIGLYMKE_1, A29491_1, A29498_1, and VPR2_MOUSE.

EXAMPLE 29: Isolation of cDNA clones Encoding Human PRO943

A consensus DNA sequence encoding PRO943 was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence was then extended using repeated cycles of BLAST and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above. The extended consensus sequence is herein designated DNA36360. Based on the DNA36360 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO943.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CGAGATGACGCCGAGCCCC-3' (SEQ ID NO:120)

reverse PCR primer 5'-CGGTTTCGACACGCGGCAGGTG-3' (SEQ ID NO:121)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA36360 sequence which had the following nucleotide sequence

hybridization probe

5'-TGCTGCTCCTGCTGCCGCCGCTGCTGCTGGGGCCTTCCCGCCGG-3' (SEQ ID NO:122)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO943 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO943 (designated herein as DNA52192-1369 [Figure 69, SEQ ID NO:118]) and the derived protein sequence for PRO943.

The entire nucleotide sequence of DNA52192-1369 is shown in Figure 69 (SEQ ID NO:118). Clone DNA52192-1369 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 150-152 and ending at the stop codon at nucleotide positions 1662-1664 (Figure 69). The predicted polypeptide precursor is 504 amino acids long (Figure 70). The full-length PRO943 protein shown in Figure 70 has an estimated molecular weight of about 54,537 daltons and a pI of about 10.04. Analysis of the full-length PRO943 sequence shown in Figure 70 (SEQ ID NO:119) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 17, a transmembrane domain from about amino acid 376 to about amino acid 396, tyrosine kinase phosphorylation sites from about amino acid 212 to about amino acid 219 and from about amino acid 329 to about amino acid 336, potential N-glycosylation sites from about amino acid 111 to about amino acid 114, from about amino acid 231 to about amino acid 234, from about amino acid 255 to about amino acid 258 and from about amino acid 293 to about amino acid 296 and an immunoglobulin and MHC protein sequence homology block from about amino acid 219 to about amino acid 236. Clone DNA52192-1369 has been deposited with ATCC on July 1, 1998 and is assigned ATCC deposit no. 203042.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 70 (SEQ ID NO:119), evidenced significant homology between the PRO943 amino acid sequence and the following Dayhoff sequences: B49151, A39752, FGR1_XENLA, S38579, RATHBFGFRB_1, TVHU2F, FGR2_MOUSE, CEK3_CHICK, P_R21080 and A27171_1.

EXAMPLE 30: Isolation of cDNA clones Encoding Human PRO1188

A consensus DNA sequence was assembled relative to other EST sequences using the program "phrap" as described in Example 1 above. This consensus sequence is designated herein as DNA45679. Based on the DNA45679 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1188.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CTGGTGCCTCAACAGGGAGCAG-3' (SEQ ID NO:125)

reverse PCR primer 5'-CCATTGTGCAGGTCAGGTCACAG-3' (SEQ ID NO:126)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA45679 sequence which had the following nucleotide sequence:

hybridization probe

5'-CTGGAGCAAGTGCTCAGCTGCCTGTGGTCAGACTGGGGTC-3' (SEQ ID NO:127)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1188 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1188 (designated herein as DNA52598-1518 [Figure 71, SEQ ID NO:123]); and the derived protein sequence for PRO1188.

The entire coding sequence of PRO1188 is shown in Figure 71 (SEQ ID NO:123). Clone DNA52598-1518 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 136-138 and an apparent stop codon at nucleotide positions 3688-3690. The predicted polypeptide precursor is 1184 amino acids long. The full-length PRO1188 protein shown in Figure 72 has an estimated molecular weight of about 132,582 Daltons and a pI of about 8.80. Additional features include: a signal peptide at about amino acids 1-31; an ATP/GTP binding site motif A (P-loop) at about amino acids 266-273; an aldehyde dehydrogenases cysteine active site at about amino acids 188-199; growth factor and cytokines receptors family signature 2 at about amino acids 153-159; and potential N-glycosylation sites at about amino acids 129-132, 132-135, 346-349, 420-423, 550-553, 631-634, 1000-1003, and 1056-1059.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 72 (SEQ ID NO:124), revealed significant homology between the PRO1188 amino acid sequence and the following Dayhoff sequences: SSU83114_1, S56015, CET21B6_4, CELT19D2_1, and TSP1_MOUSE.

Clone DNA52598-1518 has been deposited with ATCC and is assigned ATCC deposit no 203107.

EXAMPLE 31: Isolation of cDNA clones Encoding Human PRO1133

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This sequence was extended using repeated cycles of phrap. The extended consensus sequence is designated herein DNA38102. Based on the DNA38102 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1133.

PCR primers (two forward and one reverse) were synthesized:

forward PCR primer 1 5'-TCGATTATGGACGAACATGGCAGC-3' (SEQ ID NO:130);

forward PCR primer 2 5'-TTCTGAGATCCCTCATCCTC-3' (SEQ ID NO:131); and

reverse primer 5'-AGGTTTCAGGGACAGCAAGTTTGGG-3' (SEQ ID NO:132).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA38102 sequence which had the following nucleotide sequence:

hybridization probe

5'TTTGCTGGACCTCGGCTACGGAATTGGCTTCCTCTACGGACAGCTGGAT3' (SEQ ID NO:133).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with a PCR primer pair identified above. A positive library was then used to

isolate clones encoding the PRO1133 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1133 and the derived protein sequence for PRO1133.

The entire coding sequence of PRO1133 is shown in Figure 73 (SEQ ID NO:128). Clone DNA53913-1490 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 266-268 and an apparent stop codon at nucleotide positions 1580-1582 of SEQ ID NO:128. The predicted polypeptide precursor is 438 amino acids long. The signal peptide is at amino acids 1-18 of SEQ ID NO:129. EGF-like domain cysteine pattern signatures start at 315 and 385 of SEQ ID NO:129 as shown in Figure 74. Clone DNA53913-1490 has been deposited with ATCC and is assigned ATCC deposit no. 203162. The full-length PRO1133 protein shown in Figure 74 has an estimated molecular weight of about 49,260 daltons and a pI of about 6.15.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 74 (SEQ ID NO:129), revealed some sequence identity between the PRO1133 amino acid sequence and the following Dayhoff sequences (data from the database incorporated herein): AF002717_1, LMG1_HUMAN, B54665, UNC6_CAEEL, LML1_CAEEL, LMA5_MOUSE, MMU88353_1, LMA1_HUMAN, HSLN2C64_1 and AF005258_1.

EXAMPLE 32: Isolation of cDNA clones Encoding Human PRO784

An initial DNA sequence (SEQ ID NO:136), referred to herein as DNA44661 and shown in Figure 77, was identified using a yeast screen, in a human fetal lung cDNA library that preferentially represents the 5' ends of the primary cDNA clones. DNA44661 was then compared to ESTs from public databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA), using the computer program BLAST or BLAST2 [Altschul et al., *Methods in Enzymology*, 266:460-480 (1996)]. The ESTs were then clustered and assembled into a consensus DNA sequence using the computer program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained is designated herein as "DNA45463". Based on the DNA45463 consensus sequence, oligonucleotides were synthesized for use as probes to isolate a clone of the full-length coding sequence for PRO784 from a human fetal lung cDNA library.

The full length DNA53978-1443 clone shown in Figure 75 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 37-39 and ending at the stop codon found at nucleotide positions 821-823 (Figure 75; SEQ ID NO:134). The predicted polypeptide precursor (Figure 76, SEQ ID NO:135) is 228 amino acids long. PRO784 has a calculated molecular weight of approximately 25,735 Daltons and an estimated pI of approximately 5.45. PRO784 has the following features: a signal peptide at about amino acid 1 to about 15; transmembrane domains at about amino acids 68 to about 87 and at about 183 to about 204; potential N-myristoylation sites at about amino acids 15-20, 51-56, 66-60, 163-168, and 206-211; and an RNP-1 protein RNA-binding region at about amino acids 108 to about 117.

Clone DNA53978-1443 was deposited with ATCC on June 16, 1998, and is assigned ATCC deposit no. 209983.

Based on a BLAST and FastA sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO784 shows amino acid sequence identity to the following proteins: RNU42209_1, MMU91538_1, CGU91742_1, CELF55A4_6, SC22_YEAST, and F48188.

EXAMPLE 33: Isolation of cDNA Clones Encoding Human PRO783

A yeast screening assay was employed to identify cDNA clones that encoded potential secreted proteins. Use of this yeast screening assay allowed identification of a single cDNA clone, designated herein as DNA45201 (Figure 80; SEQ ID NO:139).

The DNA45201 sequence was then used to search expressed sequence tag (EST) databases for the presence of potential homologies. The EST databases included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, Univ. of Washington, Seattle, Washington). The consensus sequence obtained is herein designated as "consen01". A proprietary Genentech EST sequence was used in the consensus assembly and is herein designated as DNA14575 (Figure 81; SEQ ID NO:140).

Based on the consen01 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO783. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

PCR primers (forward and reverse) were synthesized:
forward PCR primer 5'-GACTGTATCTGAGCCCCAGACTGC-3' (SEQ ID NO:141),
forward PCR primer 5'-TCAGCAATGAGGTGCTGCTC-3' (SEQ ID NO:142), and
reverse PCR primer 5'-TGAGGAAGATGAGGGACAGGTTGG-3' (SEQ ID NO:143).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consen01 sequence which had the following nucleotide sequence:

hybridization probe

5'-TATGGAAGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCC-3' (SEQ ID NO:144).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with a PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO783 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB228). The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially

available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

5 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO783 [herein designated as DNA53996-1442] (SEQ ID NO:137) and the derived protein sequence for PRO783.

The entire nucleotide sequence of DNA53996-1442 is shown in Figure 78 (SEQ ID NO:137). Clone DNA53996-1442 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 310-312 and ending at the stop codon at nucleotide positions 1777-1779 (Figure 78). The predicted polypeptide precursor is 489 amino acids long (Figure 79). The full-length PRO783 protein shown in Figure 79 has an estimated molecular weight of about 55,219 daltons and a pI of about 8.47. Analysis of the full-length PRO783 sequence shown in Figure 79 (SEQ ID NO:138) evidences the presence of the following features: transmembrane domains located at about amino acids 23-42, 67-89, 111-135, 154-176, 194-218, 296-319, 348-370, 387-410 and 427-452; leucine zipper patterns located at about amino acids 263-283 and 399-420; a potential tyrosine kinase phosphorylation site at about amino acids 180-187; potential N-glycosylation sites at about amino acids 105 -108 and 121-124; potential cAMP- and a cGMP-dependent protein kinase phosphorylation site at about amino acids 288-291; and a region having sequence identity with bacterial rhodopsins retinal binding site protein at about amino acids 190-218.

20 An analysis of the Dayhoff database (version 35.45 SwissProt 35) shows some sequence identity between the PRO783 amino acid sequence and the following Dayhoff sequences: YNC2_CAEEL, D64048, ATAC002332_3F4P9.3, NY2R_SHEEP, and VSH_MUMPA.

Clone DNA53996-1442 was deposited with the ATCC on June 2, 1998, and is assigned ATCC deposit no. 209921.

25 EXAMPLE 34: Isolation of cDNA Clones Encoding Human PRO820

An expressed sequence tag (EST) DNA database (Merck/Wash. U) was searched and an EST designated EST no. AA504080, Merck clone 825136, was identified (library 312, human B-cell tonsil). Homology searches revealed that this EST showed sequence identity with low affinity immunoglobulin gamma Fc receptor II. DNA sequencing gave the full-length DNA sequence for PRO820 and the derived protein sequence for PRO820.

35 The entire nucleotide sequence of DNA56041-1416 is shown in Figure 82 (SEQ ID NO:145). Clone DNA56041-1416 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 115-117 and ending at the stop codon at nucleotide positions 487-489 (Figure 82). The predicted polypeptide precursor is 124 amino acids long (Figure 83). The full-length PRO820 protein shown in Figure 83 has an estimated molecular weight of about 14,080 daltons and a pI of about 7.48. Clone DNA56041-1416 has been deposited with ATCC. Regarding the sequence, it is understood that the deposited clone contains the

correct sequence, and the sequences provided herein are based on known sequencing techniques.

Still analyzing the amino acid sequence of SEQ ID NO:146, the putative signal peptide is at about amino acids 1-15 of SEQ ID NO:146. Protein kinase C phosphorylation sites are at about amino acids 20-22 and 43-45 of SEQ ID NO:146. An N-myristoylation site is at about amino acids 89-94 of SEQ ID NO:146. An immunoglobulin and major histocompatibility complex domain is at about amino acids 83-90 of SEQ ID NO:146. The corresponding nucleotides can be routinely determined given the sequences provided herein.

EXAMPLE 35: Isolation of cDNA Clones Encoding Human PRO1080

A consensus DNA sequence was assembled relative to other EST sequences using phrap and was extended using repeated cycles of BLAST and phrap so as to extend the consensus sequence as far as possible using the sources of the EST sequences as described in Example 1 above. The consensus sequence is designated herein as DNA52640. An EST proprietary to Genentech was employed in the consensus assembly and is herein designated as DNA36527 (Figure 86; SEQ ID NO:149).

In light of an observed sequence homology between the DNA36527 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 526423, the Merck EST clone 526423 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 84 and is herein designated as DNA56047-1456.

The entire nucleotide sequence of DNA56047-1456 is shown in Figure 84 (SEQ ID NO:147). Clone DNA56047-1456 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 159-161 and ending at the stop codon at nucleotide positions 1233-1235 of SEQ ID NO:147 (Figure 84). The predicted polypeptide precursor is 358 amino acids long (Figure 85). The full-length PRO1080 protein shown in Figure 85 has an estimated molecular weight of about 40,514 daltons and a pI of about 6.08. Clone DNA56047-1456 has been deposited with ATCC on June 9, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Also shown in Figure 85 are the approximate locations of the signal peptide, cell attachment site, Nt-DnaJ domain signature, region having sequence identity with Nt-DnaJ domain proteins, and N-glycosylation sites. The corresponding nucleic acids of these amino acid sequences and others provided herein can be routinely determined by the information provided herein.

EXAMPLE 36: Isolation of cDNA Clones Encoding Human PRO1079

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above, and is herein designated DNA52714. Based on information provided by the assembly, the clone for Merck EST no. HO6898 was obtained and sequenced, thereby giving the nucleotide sequence designated herein as DNA56050-1455. The entire nucleotide sequence of DNA56050-1455 is shown in Figure 87 (SEQ ID NO:150). Clone DNA56050-1455 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 183-185 and ending at the stop codon at nucleotide positions 861-863 (Figure 87). The predicted polypeptide precursor is 226 amino acids long (Figure 88). The full-

length PRO1079 protein shown in Figure 88 has an estimated molecular weight of about 24,611 Daltons and a pI of about 4.85. Analysis of the full-length PRO1079 sequence shown in Figure 88 (SEQ ID NO:3) evidences the presence of the following features: a signal peptide at about amino acid 1-29; potential N-myristoylation sites at about amino acids 10-15, and 51-56; homology to photosystem I psaG and psaK proteins at about amino acids 2 to 20; and homology to prolyl endopeptidase family serine proteins at about amino acids 150 to 163.

Analysis of the amino acid sequence of the full-length PRO1079 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced some sequence identity between the PRO1079 amino acid sequence and the following Dayhoff sequences: CEK10C3_4, MMU50734_1, D69503, AF051149_1, and VSMP_CVMS.

Clone UNQ536 (DNA56050-1455) was deposited with the ATCC on June 22, 1998, and is assigned ATCC deposit no. 203011.

EXAMPLE 37: Isolation of cDNA clones Encoding Human PRO793

A cDNA clone (DNA56110-1437) encoding a native human PRO793 polypeptide was identified by a yeast screen, in a human skin tumor cDNA library that preferentially represents the 5' ends of the primary cDNA clones. The yeast screen employed identified a single EST clone designated herein as DNA50177 (Figure 91; SEQ ID NO:154). The DNA50177 sequence was then compared to various EST databases including public EST databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify homologous EST sequences. The comparison was performed using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)]. Those comparisons resulting in a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). This consensus sequence is herein designated DNA50972.

In light of an observed sequence homology between the DNA50972 consensus sequence and an EST sequence encompassed within the Merck EST clone no. N33874, the Merck EST clone N33874 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 89 and is herein designated as DNA56110-1437.

The full-length DNA56110-1437 clone shown in Figure 89 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 77-79 and ending at the stop codon at nucleotide positions 491-493 (Figure 89). The predicted polypeptide precursor is 138 amino acids long (Figure 90). The full-length PRO793 protein shown in Figure 90 has an estimated molecular weight of about 15,426 daltons and a pI of about 10.67. Analysis of the full-length PRO793 sequence shown in Figure 90 (SEQ ID NO:153) evidences the presence of the following: transmembrane domains from about amino acid 12 to about amino acid 30, from about amino acid 33 to about amino acid 52, from about amino acid 69 to about amino acid 89 and from about amino acid 93 to about amino acid 109, potential N-myristoylation sites from about amino acid 11 to about amino acid 16, from about amino acid 51 to about amino acid 56 and from about amino acid 116

to about amino acid 121 and an amino acid sequence block having homology to an aminoacyl-transfer RNA synthetase class-II protein from about amino acid 49 to about amino acid 59. Clone DNA56110-1437 has been deposited with ATCC on August 11, 1998 and is assigned ATCC deposit no. 203113.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 90 (SEQ ID NO:153), evidenced certain homology between the PRO793 amino acid sequence and the following Dayhoff sequences: S47453, AF015193_12, MTEHGNS9_2, E64030, H69784, D64995, CD53_MOUSE, GEN8006, AE001138_7 and COX2_STRPU.

EXAMPLE 38: Isolation of cDNA Clones Encoding Human PRO1016

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. The consensus sequence obtained is herein designated DNA53502.

In light of an observed sequence homology between the DNA53502 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 38680, the Merck EST clone 38680 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 92.

The entire nucleotide sequence of DNA56113-1378 is shown in Figure 92 (SEQ ID NO:155). Clone DNA56113-1378 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 168-170 and ending at the stop codon at nucleotide positions 1302-1304 (Figure 92). The predicted polypeptide precursor is 378 amino acids long (Figure 93). The full-length PRO1016 protein shown in Figure 93 has an estimated molecular weight of about 44,021 daltons and a pI of about 9.07. Clone DNA56113-1378 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO1016 polypeptide suggests that portions of it possess sequence identity with acyltransferase, thereby indicating that PRO1016 may be a novel acyltransferase.

Still analyzing the amino acid sequence of SEQ ID NO:156, the putative signal peptide is at about amino acids 1-18 of SEQ ID NO:156. The transmembrane domain(s) are at about amino acids 332-352 and 305-330 of SEQ ID NO:156. The fructose-bisphosphate aldolase class-II protein homology sequence is at about amino acids 73-90 of SEQ ID NO:156. The extradiol ring-cleavage dioxygenase protein is at about amino acids 252-275 of SEQ ID NO:156. The corresponding nucleotides can be routinely determined given the sequences provided herein.

The specific Dayhoff database designation names of sequences to which PRO1016 has sequence identity with include the following: S52645, P_R59712, P_R99249, P_R59713, BNAGPATRF_1, CELT05H4_15 and CELZK40_1.

EXAMPLE 39: Isolation of cDNA Encoding Human PRO1013

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described

in Example 1 above. The consensus DNA sequence was then extended using repeated cycles of BLAST and phrap to extend the consensus sequence as far as possible using the sources of EST sequences.

In light of an observed sequence homology between the consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3107695, the Incyte EST clone 3107695 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 94 and is herein designated as DNA56410-1414.

The entire nucleotide sequence of DNA56410-1414 is shown in Figure 94 (SEQ ID NO:157). Clone DNA56410-1414 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 17-19 and ending at the stop codon at nucleotide positions 1244-1246 (Figure 94). The predicted polypeptide precursor is 409 amino acids long (Figure 95). The full-length PRO1013 protein shown in Figure 95 has an estimated molecular weight of about 46,662 daltons and a pI of about 7.18. Clone DNA56410-1414 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Still analyzing the amino acid sequence of SEQ ID NO:158, the putative signal peptide is at about amino acids 1-19 of SEQ ID NO:158. N-glycosylation sites are at about amino acids 75-78 and 322-325 of SEQ ID NO:158. An N-myristoylation site is at about amino acids 184-189 of SEQ ID NO:158. A growth factor and cytokine receptor family domain is at about amino acids 134-149 of SEQ ID NO:158. The corresponding nucleotides can be routinely determined given the sequences provided herein.

Blast analysis showed some sequence identity with other proteins. Specifically, PRO1013 has some sequence identity with at least the Dayhoff sequences designated: D63877_1; MHU22019_1, AE000730_10, and AF019079_1.

EXAMPLE 40: Isolation of cDNA Clones Encoding Human PRO937

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. That consensus sequence is herein designated DNA49651. Based on the DNA49651 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO937.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CTCCGTGGTAAACCCACAGCCC-3' (SEQ ID NO:161); and

reverse PCR primer 5'-TCACATCGATGGGATCCATGACCG-3' (SEQ ID NO:162).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA48651 sequence which had the following nucleotide sequence:

hybridization probe

5'-GGTCTCGTGACTGTGAAGCCATGTTACAACACTACTGCTCAAACATCATGAG-3' (SEQ ID NO:163).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO937 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO937 [herein designated as DNA56436-1448] (SEQ ID NO:159) and the derived protein sequence for PRO937.

The entire nucleotide sequence of DNA56436-1448 is shown in Figure 96 (SEQ ID NO:159). It contains a single open reading frame having an apparent translational initiation site at nucleotide positions 499-501 and ending at the stop codon found at nucleotide positions 2167-2169 (Figure 96, SEQ ID NO:159). The predicted polypeptide precursor is 556 amino acids long, has a calculated molecular weight of approximately 62,412 daltons and an estimated pI of approximately 6.62. Analysis of the full-length PRO937 sequence shown in Figure 97 (SEQ ID NO:160) evidences the presence of the following features: signal peptide at about amino acids 1-22; ATP/GTP-binding site motif A (P-loop) at about amino acids 515-523; a potential N-glycosylation site at about amino acids 514-517; and sites of glypican homology at about amino acids 54-74, 106-156, 238-279, 309-345, 423-459, and 468-505.

Clone DNA56436-1448 has been deposited with ATCC on May 27, 1998, and is assigned ATCC deposit no. 209902.

Analysis of the amino acid sequence of the full-length PRO937 polypeptide suggests that it possesses significant sequence similarity to glypican proteins, thereby indicating that PRO937 may be a novel glypican protein. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO937 amino acid sequence and the following Dayhoff sequences: GPCCK_MOUSE, GPC2_RAT, GPC5_HUMAN, GPC3_HUMAN, P_R30168, CEC03H12_2, GEN13820, HS119E23_1, HDAC_DROME, and AF017637_1.

EXAMPLE 41: Isolation of cDNA clones Encoding Human PRO842

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence designated herein as Incyte EST cluster sequence no. 69572. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA54230.

In light of an observed sequence homology between the consensus sequence and an EST sequence encompassed within the Merck EST clone no. AA477092, the Merck EST clone AA477092 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 98 and is herein designated as DNA56855-1447.

The full length clone shown in Figure 98 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 153-155 and ending at the stop codon found at nucleotide positions 510-512 (Figure 98; SEQ ID NO:164). The predicted polypeptide precursor (Figure 99, SEQ ID NO:165) is 119 amino acids long. PRO842 has a calculated molecular weight of approximately 13,819 Daltons and an estimated pI of approximately 11.16. Other features of PRO842 include a signal peptide at about amino acids 1-22, a potential protein kinase C phosphorylation site at about amino acids 39-41 and two potential N-myristoylation sites at about amino acids 27-32 and about amino acids 46-51.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 98 (SEQ ID NO:164), evidenced some homology between the PRO842 amino acid sequence and the following Dayhoff sequences: CEZK131_11, P_R80843, RAT5HT2X_1, S81882_1, A60912, MCU60315_137MC137L, U93422_1, p_P91996, U93462_1, and ZN18_HUMAN.

Clone DNA56855-1447 was deposited with the ATCC on June 23, 1998, and is assigned ATCC deposit no. 203004.

EXAMPLE 42: Isolation of cDNA clones Encoding Human PRO839

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte LIFESEQ® database, designated Incyte EST Cluster No. 24479. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55709.

In light of an observed sequence homology between the DNA55709 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 754525, the Merck EST clone 754525 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 100 and is herein designated as DNA56859-1445.

The full length clone shown in Figure 100 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 2-4 and ending at the stop codon found at nucleotide positions 263-265 (Figure 100; SEQ ID NO:166). The predicted polypeptide precursor (Figure 101, SEQ ID NO:167) is 87 amino acids long. PRO839 has a calculated molecular weight of approximately 9,719 Daltons and an estimated pI of approximately 4.67. Other features of PRO839 include a signal peptide at about amino acids 1-23, potential protein kinase C phosphorylation sites at about amino acids 37-39 and about amino acids 85-87, a potential casein kinase II phosphorylation site at about amino acids 37-40, sequence identity with ribonucleotide reductase large subunit protein at about amino acids 50-60, and sequence identity with

eukaryotic RNA-binding region RNP-1 proteins at about amino acids 70-79.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 101 (SEQ ID NO:167), evidenced some homology between the PRO839 amino acid sequence and the following Dayhoff sequences: CD14_MOUSE, XPR6_YARLI, HS714385_1, S49783, BB19_RABIT, GVPH-HALME, AB003135_1, P_R85453, LUU27081_2, and TP2B_MOUSE.

Clone DNA56859-1445 was deposited with the ATCC on June 23, 1998, and is assigned ATCC deposit no.209019.

EXAMPLE 43: Isolation of cDNA Clones Encoding Human PRO1180

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence (Incyte EST cluster sequence no. 14732). The Incyte EST cluster sequence no. 14732 sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55711.

In light of an observed sequence homology between the DNA55711 consensus sequence and an EST sequence encompassed within the Merck EST clone no. T60981, the Merck EST clone T60981 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 102 and is herein designated DNA56860-1510.

The full length clone shown in Figure 102 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 78-80 and ending at the stop codon found at nucleotide positions 909-911 (Figure 102; SEQ ID NO:168). The predicted polypeptide precursor is 277 amino acids long, has a calculated molecular weight of approximately 31,416 daltons and an estimated pI of approximately 8.88. Analysis of the full-length PRO1180 sequence shown in Figure 103 (SEQ ID NO:169) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 23, a leucine zipper pattern sequence from about amino acid 10 to about amino acid 31, and potential N-myristylation sited from about amino acid 64 to about amino acid 69, from about amino acid 78 to about amino acid 83, from about amino acid 80 to about amino acid 85, from about amino acid 91 to about amino acid 96 and from about amino acid 201 to about amino acid 206. Clone DNA56860-1510 has been deposited with the ATCC on June 9, 1998 and is assigned ATCC deposit no. 209952.

Analysis of the amino acid sequence of the full-length PRO1180 polypeptide suggests that it possesses sequence similarity to the methyltransferase family of proteins. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some degree of homology between the PRO1180 amino acid

sequence and the following Dayhoff sequences, MTCI65_14, D69267, YH09_YEAST, BIOC_SERMA, ATAC00448415T1D16.16, SHGCPIR_18, SPBC3B9_4, AB009504_14, P_W17977 and A69952.

EXAMPLE 44: Isolation of cDNA clones Encoding Human PRO1134

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 7511. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55725. Two proprietary Genentech EST sequences were employed in the assembly and are shown in Figure 106 (SEQ ID NO:172) and Figure 107 (SEQ ID NO:173).

In light of an observed sequence homology between the DNA55725 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H94897, the Merck EST clone H94897 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 106 and is herein designated as DNA56865-1491.

Clone DNA56865-1491 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 153-155 and ending at the stop codon at nucleotide positions 1266-1268 (Figure 104). The predicted polypeptide precursor is 371 amino acids long (Figure 105). The full-length PRO1134 protein shown in Figure 105 has an estimated molecular weight of about 41,935 daltons and a pI of about 9.58. Analysis of the full-length PRO1134 sequence shown in Figure 105 (SEQ ID NO:171) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 23, potential N-glycosylation sites from about amino acid 103 to about amino acid 106, from about amino acid 249 to about amino acid 252 and from about amino acid 257 to about amino acid 260, and an amino acid block having homology to tyrosinase CuA-binding region proteins from about amino acid 280 to about amino acid 306. Clone DNA56865-1491 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203022.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 105 (SEQ ID NO:171), evidenced significant homology between the PRO1134 amino acid sequence and the following Dayhoff sequences: F20P5_18, AC002396_10, S47847, C64146, GSPA_BACSU, P_W10564, RFAI_ECOLI, Y258_HAEIN, RFAJ_SALTY and P_R32985.

EXAMPLE 45: Isolation of cDNA clones Encoding Human PRO830

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incytedatabase, designated 20251. This EST cluster sequence was then compared

to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., *Methods in Enzymology* 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55733.

In light of an observed sequence homology between the DNA55733 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H78534, the Merck EST clone H78534 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 108 and is herein designated as DNA56866-1342.

Clone DNA56866-1342 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 154-156 and ending at the stop codon at nucleotide positions 415-417 (Figure 108). The predicted polypeptide precursor is 87 amino acids long (Figure 109). The full-length PRO830 protein shown in Figure 109 has an estimated molecular weight of about 9,272 daltons and a pI of about 9.19. Analysis of the full-length PRO830 sequence shown in Figure 109 (SEQ ID NO:175) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 33, potential N-myristoylation sites from about amino acid 2 to about amino acid 7 and from about amino acid 8 to about amino acid 13 and a thioredoxin family of proteins homology block from about amino acid 23 to about amino acid 39. Clone UNQ470 (DNA56866-1342) has been deposited with ATCC on June 22, 1998 and is assigned ATCC deposit no. 203023.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 109 (SEQ ID NO:175), evidenced significant homology between the PRO830 amino acid sequence and the following Dayhoff sequences: HSU88154_1, HSU88153_1, SAPKSGENE_1, HPU31791_5, GGCNOT2_1, CPU91421_1, CHKESTPC09_1, PQ0769, U97553_79 and B60095.

EXAMPLE 46: Isolation of cDNA clones Encoding Human PRO1115

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the LIFESEQ® database, designated Incyte EST cluster sequence no. 165008. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., *Methods in Enzymology* 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55726.

In light of an observed sequence homology between the DNA55726 consensus sequence and an EST sequence encompassed within the Merck EST clone no. R75784, the Merck EST clone R75784 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 111 and is herein designated as DNA56868-1478.

The full length clone shown in Figure 110 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 189-191 and ending at the stop codon found at nucleotide positions 1524-1526 (Figure 110; SEQ ID NO:176). The predicted polypeptide precursor (Figure 111, SEQ ID NO:177) is 445 amino acids long. PRO1115 has a calculated molecular weight of approximately 50,533 Daltons and an estimated pI of approximately 8.26. Additional features include a signal peptide at about amino acids 1-20; potential N-glycosylation sites at about amino acids 204-207, 295-298, and 313-316; and putative transmembrane domains at about amino acids 35-54, 75-97, 126-146, 185-204, 333-350, and 353-371.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 111 (SEQ ID NO:177), evidenced some amino acid sequence identity between the PRO1115 amino acid sequence and the following Dayhoff sequences: AF053947_79, S73698, CEC47A10_4, CCOMTNS5G_1, HS4LMP2AC_1, LMP2_EBV, PA24_MOUSE, HCU33331_7, P-W05508, and AF002273_1.

Clone DNA56868-1478 was deposited with the ATCC on June 23, 1998 and is assigned ATCC deposit no. 203024..

EXAMPLE 47: Isolation of cDNA clones Encoding Human PRO1277

A consensus DNA sequence was assembled relative to other ESTs using repeated cycles of BLAST and the program "phrap" as described in Example 1 above. One or more of the ESTs from the assembly was derived from diseased coronary artery tissue. The consensus sequence obtained is designated herein as "DNA49434".

In light of an observed sequence homology between the DNA49434 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3042605, the Incyte EST clone 3042605 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 112 (SEQ ID NO:178).

Clone DNA56869-1545 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 188-190, and an apparent stop codon at nucleotide positions 2222-2224 (Figure 112). The predicted polypeptide precursor is 678 amino acids long (Figure 113). The full-length PRO1277 protein shown in Figure 113 has an estimated molecular weight of about 73,930 daltons and a pI of about 9.48. Additional features include a signal peptide at about amino acids 1-26; a transmembrane domain at about amino acids 181-200, and potential N-glycosylation sites at about amino acids 390-393 and 520-523.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 113 (SEQ ID NO:179), revealed significant homology between the PRO1277 amino acid sequence and Dayhoff sequence no AF012252_1. Homology was also found between the PRO1277 amino acid sequence and the following Dayhoff sequences: AF006740_1,

CA36_HUMAN, HSU1_1, HUMCOL7A1X_1, CA17_HUMAN, MMZ78163_1, CAMA_CHICK, HSU69263_1, YNX3_CAEEL, and MMRNAM3_1.

Clone DNA56869-1545 has been deposited with ATCC and is assigned ATCC deposit no. 203161.

EXAMPLE 48: Isolation of cDNA Clones Encoding Human PRO1135

5 A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA52767. Based on the DNA52767 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1135.

10 In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with PCR primer pairs prepared based upon the DNA52767 sequence. A positive library was then used to isolate clones encoding the PRO1135 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human coronary artery smooth muscle tissue (LIB309). The cDNA libraries used to isolate the cDNA clones were constructed
15 by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

20 DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1135 [herein designated as DNA56870-1492] (SEQ ID NO:180) and the derived protein sequence for PRO1135.

The entire nucleotide sequence of DNA56870-1492 is shown in Figure 114 (SEQ ID NO:180). Clone DNA56870-1492 contains a single open reading frame with an apparent translational initiation site at nucleotide
25 positions 62-64 and ending at the stop codon at nucleotide positions 1685-1687 (Figure 114). The predicted polypeptide precursor is 541 amino acids long (Figure 115). The full-length PRO1135 protein shown in Figure 115 has an estimated molecular weight of about 60,335 daltons and a pI of about 5.26. Analysis of the full-length PRO1135 sequence shown in Figure 115 (SEQ ID NO:181) evidences the presence of the following:
30 a signal peptide from about amino acid 1 to about amino acid 21, potential N-glycosylation sites from about amino acid 53 to about amino acid 56, from about amino acid 75 to about amino acid 78, from about amino acid 252 to about amino acid 255 and from about amino acid 413 to about amino acid 416 and an amino acid block having homology to glycosyl hydrolase family 35 proteins from about amino acid 399 to about amino acid 414. Clone DNA56870-1492 has been deposited with ATCC on June 2, 1998 and is assigned ATCC deposit no. 209925.

35 Analysis of the amino acid sequence of the full-length PRO1135 polypeptide suggests that it possesses significant sequence similarity to the alpha 1,2-mannosidase protein, thereby indicating that PRO1135 may be a novel mannosidase. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35)

evidenced significant homology between the PRO1135 amino acid sequence and the following Dayhoff sequences, DMC86E4_5, D86967_1, SPAC23A1_4, YH04_YEAST, B54408, SSMAN9MAN_1, CEZC410_4, S61631 and MSU14190_1.

EXAMPLE 49: Isolation of cDNA Clones Encoding Human PRO1114

A cDNA sequence isolated in the amylase screen described in Example 2 above was found, by the WU-BLAST-2 sequence alignment computer program, to have certain sequence identity to other known interferon receptors. This cDNA sequence is herein designated DNA48466 and is shown in Figure 118 (SEQ ID NO:184). Based on the sequence identity, probes were generated from the sequence of the DNA48466 molecule and used to screen a human breast carcinoma library (LIB135) prepared as described in paragraph 1 of Example 2 above. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, 253:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp.

The oligonucleotide probes employed were as follows:

forward PCR primer 5'-AGGCTTCGCTGCGACTAGACCTC-3' (SEQ ID NO:185)

reverse PCR primer 5'-CCAGGTCGGGTAAGGATGGTTGAG-3' (SEQ ID NO:186)

hybridization probe

5'-TTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC-3' (SEQ ID NO:187)

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 250-252, and a stop signal at nucleotide positions 1183-1185 (Figure 116, SEQ ID NO:182). The predicted polypeptide precursor is 311 amino acids long, has a calculated molecular weight of approximately 35,076 daltons and an estimated pI of approximately 5.04. Analysis of the full-length PRO1114 interferon receptor sequence shown in Figure 117 (SEQ ID NO:183) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 29, a transmembrane domain from about amino acid 230 to about amino acid 255, potential N-glycosylation sites from about amino acid 40 to about amino acid 43 and from about amino acid 134 to about amino acid 137, an amino acid sequence block having homology to tissue factor proteins from about amino acid 92 to about amino acid 119 and an amino acid sequence block having homology to integrin alpha chain proteins from about amino acid 232 to about amino acid 262. Clone DNA57033-1403 has been deposited with ATCC on May 27, 1998 and is assigned ATCC deposit no. 209905.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 117 (SEQ ID NO:183), evidenced significant homology between the PRO1114 interferon receptor amino acid sequence and the following Dayhoff sequences: G01418, INR1_MOUSE, P_R71035, INGS_HUMAN, A26595_1, A26593_1, I56215 and TF_HUMAN.

EXAMPLE 50: Isolation of cDNA Clones Encoding Human PRO828

A consensus DNA sequence was identified using the method described in Example 1 above. This

consensus sequence is herein designated DNA35717. Based on the DNA35717 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO828.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GCAGGACTTCTACGACTTCAAGGC-3' (SEQ ID NO:190); and

reverse PCR primer 5'-AGTCTGGGCCAGGTACTTGAAGGC-3' (SEQ ID NO:191).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35717 sequence which had the following nucleotide sequence:

hybridization probe

5'-CAACATCCGGGGCAAACCTGGTGTCTGCTGGAGAAGTACCGCGGATCGGTGT-3' (SEQ ID NO:192)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO828 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal lung tissue (LIB25).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO828 [herein designated as DNA57037-1444] (SEQ ID NO:188) and the derived protein sequence for PRO828.

The entire nucleotide sequence of DNA57037-1444 is shown in Figure 119 (SEQ ID NO:188). Clone DNA57037-1444 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 34-36 and ending at the stop codon at nucleotide positions 595-597 (Figure 119). The predicted polypeptide precursor is 187 amino acids long (Figure 120). The full-length PRO828 protein shown in Figure 120 has an estimated molecular weight of about 20,996 daltons and a pI of about 8.62. Analysis of the full-length PRO828 sequence shown in Figure 120 (SEQ ID NO:189) evidences the presence of the following: a signal peptide at about amino acids 1- 21; sequences identity to glutathione peroxidases signature 2 at about amino acids 82-89; sequence identity to glutathione peroxidases selenocysteine proteins at about amino acids 35-60, 63-100, 107-134, and 138-159. Clone DNA57037-1444 has been deposited with ATCC on May 27, 1998, and is assigned ATCC deposit no. 209903.

Analysis of the amino acid sequence of the full-length PRO828 polypeptide suggests that it possesses significant sequence similarity to glutathione peroxidases, thereby indicating that PRO828 may be a novel peroxidase enzyme. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced sequence identity between the PRO828 amino acid sequence and the following Dayhoff sequences: AF053311_1, CELT09A12_2, AC004151_3, BTUE_ECOLI, CER05H10_3, P_P80918, PWU88907_1, and P_W22308.

EXAMPLE 51: Isolation of cDNA clones Encoding Human PRO1009

A cDNA clone (DNA57129-1413) encoding a native human PRO1009 polypeptide was identified by the use of a yeast screen, in a human SK-Lu-1 adenocarcinoma cell line cDNA library that preferentially represents the 5' ends of the primary cDNA clones. First SEQ ID NO:195 (Figure 123) was identified, which

was extended by alignments to other EST sequences to form a consensus sequence. Oligonucleotide probes based upon the consensus sequence were synthesized and used to screen the cDNA library which gave rise to the full-length DNA57129-1413 clone.

The full length DNA57129-1413 clone shown in Figure 121 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 41-43 and ending at the stop codon found at nucleotide positions 1886-1888 (Figure 121; SEQ ID NO:193). The predicted polypeptide precursor (Figure 122, SEQ ID NO:194) is 615 amino acids long. Figure 122 also shows the approximate locations of the signal sequence, transmembrane domains, myristoylation sites, a glycosylation site and an AMP-binding domain. PRO1009 has a calculated molecular weight of approximately 68,125 daltons and an estimated pI of approximately 7.82. Clone DNA57129-1413 has been deposited with ATCC and is assigned ATCC deposit no. 209977. It is understood that the deposited clone has the actual and correct sequence and that the representations herein may have minor, normal sequencing errors.

Based on a WU-BLAST-2 sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO1009 shows amino acid sequence identity to at least the following proteins which were designated in a Dayhoff database as follows: F69893, CEF28F8_2, BSY13917_7, BSY13917_7, D69187, D69649, XCRPFB_1, E64928, YDID_ECOLI, BNACSF8_1 and RPU75363_2.

EXAMPLE 52: Isolation of cDNA Clones Encoding Human PRO1007

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated as DNA40671.

In light of an observed sequence homology between the DNA40671 consensus sequence and an EST sequence encompassed within the Merck EST clone no. T70513, the Merck EST clone T70513 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 124.

The entire nucleotide sequence of DNA57690-1374 is shown in Figure 124 (SEQ ID NO:196). Clone DNA57690-1374 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 16-18 and ending at the stop codon at nucleotide positions 1054-1056 (Figure 124). The predicted polypeptide precursor is 346 amino acids long (Figure 125). The full-length PRO1007 protein shown in Figure 125 has an estimated molecular weight of about 35,971 daltons and a pI of about 8.17. Clone DNA57690-1374 has been deposited with the ATCC on June 9, 1998. Regarding the sequence, it is understood that the deposited clone contains the actual sequence, and the sequences provided herein are based on known sequencing techniques. The representative figures herein show the representative numbering.

Analysis of the amino acid sequence of the full-length PRO1007 polypeptide suggests that portions of it possess sequence identity to MAGPIAP, thereby indicating that PRO1007 may be a novel member of the family to which MAGPIAP belongs.

Still analyzing the amino acid sequence of SEQ ID NO:197, the putative signal peptide is at about amino acids 1-30 of SEQ ID NO:197. The transmembrane domain is at amino acids 325-346 of SEQ ID NO:197. N-glycosylation sites are at about amino acids 118-121, 129-132, 163-166, 176-179, 183-186 and

227-130 of SEQ ID NO:197. Ly-6/u-Par domain protein homology is at about amino acids 17-36 and 209-222 of SEQ ID NO:197. The corresponding nucleotides of the amino acids presented herein can be routinely determined given the sequences provided herein.

EXAMPLE 53: Isolation of cDNA clones Encoding Human PRO1056

5 Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated herein as 6425. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2
10 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55736.

15 In light of an observed sequence homology between the DNA55736 consensus sequence and an EST sequence encompassed within the Merck EST clone no. R88049, the Merck EST clone R88049 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 126 and is herein designated as DNA57693-1424.

20 Clone DNA57693-1424 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 56-58 and ending at the stop codon at nucleotide positions 416-418 (Figure 126). The predicted polypeptide precursor is 120 amino acids long (Figure 127). The full-length PRO1056 protein shown in Figure 127 has an estimated molecular weight of about 13,345 daltons and a pI of about 5.18. Analysis of the full-length PRO1056 sequence shown in Figure 127 (SEQ ID NO:199) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 18, a transmembrane domain from about amino acid 39 to about amino acid 58, a potential N-glycosylation site from about amino acid 86
25 to about amino acid 89, protein kinase C phosphorylation sites from about amino acid 36 to about amino acid 38 and from about amino acid 58 to about amino acid 60, a tyrosine kinase phosphorylation site from about amino acid 25 to about amino acid 32 and an amino acid sequence block having homology to channel forming colicin proteins from about amino acid 24 to about amino acid 56. Clone DNA57693-1424 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203008.

30 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 127 (SEQ ID NO:199), evidenced significant homology between the PRO1056 amino acid sequence and the following Dayhoff sequences: PLM_HUMAN, A40533, ATNG_HUMAN, A55571, ATNG_SHEEP, S31524, GEN13025, RIC_MOUSE, A48678 and A10871_1.

EXAMPLE 54: Isolation of cDNA clones Encoding Human PRO826

35 Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST

cluster sequence from the Incyte database, designated 47283. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56000.

In light of an observed sequence homology between the DNA56000 consensus sequence and an EST sequence encompassed within the Merck EST clone no. W69233, the Merck EST clone W69233 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 128 and is herein designated as DNA57694-1341.

Clone DNA57694-1341 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 13-15 and ending at the stop codon at nucleotide positions 310-312 (Figure 128). The predicted polypeptide precursor is 99 amino acids long (Figure 129). The full-length PRO826 protein shown in Figure 129 has an estimated molecular weight of about 11,050 daltons and a pI of about 7.47. Analysis of the full-length PRO826 sequence shown in Figure 129 (SEQ ID NO:201) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 22, potential N-myristoylation sites from about amino acid 22 to about amino acid 27 and from about amino acid 90 to about amino acid 95 and an amino acid sequence block having homology to peroxidase from about amino acid 16 to about amino acid 48. Clone DNA57694-1341 has been deposited with ATCC on June 22, 1998 and is assigned ATCC deposit no. 203017.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 129 (SEQ ID NO:201), evidenced significant homology between the PRO826 amino acid sequence and the following Dayhoff sequences: CCU12315_1, SCU96108_6, CELF39F10_4 and HELT_HELHO.

EXAMPLE 55: Isolation of cDNA clones Encoding Human PRO819

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 49605. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56015.

In light of an observed sequence homology between the DNA56015 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H65785, the Merck EST clone H65785 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 130 and is herein designated as DNA57695-1340.

Clone DNA57695-1340 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 46-48 and ending at the stop codon at nucleotide positions 202-204 (Figure 130). The predicted polypeptide precursor is 52 amino acids long (Figure 131). The full-length PRO819 protein shown in Figure 131 has an estimated molecular weight of about 5,216 daltons and a pI of about 4.67. Analysis of the full-length PRO819 sequence shown in Figure 131 (SEQ ID NO:203) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 24, a potential N-myristoylation site from about amino acid 2 to about amino acid 7 and a region having homology to immunoglobulin light chain from about amino acid 5 to about amino acid 33. Clone DNA57695-1340 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203006.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 131 (SEQ ID NO:203), evidenced significant homology between the PRO819 amino acid sequence and the following Dayhoff sequences: HSU03899_1, HUMIGLITEB_1, VG28_HSVSA, AF031522_1, PAD1_YEAST and AF045484_1.

EXAMPLE 56: Isolation of cDNA Clones Encoding Human PRO1006

An initial candidate sequence from Incyte cluster sequence no. 45748 was identified using the signal algorithm process described in Example 3 above. This sequence was then aligned with a variety of public and Incyte EST sequences and a consensus sequence designated herein as DNA56036 was derived therefrom.

In light of an observed sequence homology between the DNA56036 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 489737, the Merck EST clone 489737 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 132.

The entire nucleotide sequence of DNA57699-1412 is shown in Figure 132 (SEQ ID NO:204). Clone DNA57699-1412 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 28-30 and ending at the stop codon at nucleotide positions 1204-1206 (Figure 132). The predicted polypeptide precursor is 392 amino acids long (Figure 133). The full-length PRO1006 protein shown in Figure 133 has an estimated molecular weight of about 46,189 daltons and a pI of about 9.04. Clone DNA57699-1412 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analyzing the amino acid sequence of SEQ ID NO:205, the putative signal peptide is at about amino acids 1-23 of SEQ ID NO:205. The N-glycosylation sites are at about amino acids 40-43, 53-56, 204-207 and 373-376 of SEQ ID NO:205. An N-myristoylation site is at about amino acids 273-278 of SEQ ID NO:205.

The corresponding nucleotides of these amino acid regions and others can be routinely determined given the sequences provided herein.

EXAMPLE 57: Isolation of cDNA Clones Encoding Human PRO1112

Use of the signal sequence algorithm described in Example 3 above allowed identification of a specific EST cluster sequence. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56018.

In light of an observed sequence homology between the DNA56018 consensus sequence and an EST sequence encompassed within the Merck EST clone no. AA223546, the Merck EST clone AA223546 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 134 and is herein designated as DNA57702-1476.

The entire nucleotide sequence of DNA57702-1476 is shown in Figure 134 (SEQ ID NO:206). Clone DNA57702-1476 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 20-22 and ending at the stop codon at nucleotide positions 806-808 of SEQ ID NO:206 (Figure 134). The predicted polypeptide precursor is 262 amino acids long (Figure 135). The full-length PRO1112 protein shown in Figure 135 has an estimated molecular weight of about 29,379 daltons and a pI of about 8.93. Figure 135 also shows the approximate locations of the signal peptide and transmembrane domains. Clone DNA57702-1476 has been deposited with the ATCC on June 9, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO1112 polypeptide suggests that it possesses some sequence similarity to other proteins. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some sequence identity between the PRO1112 amino acid sequence and at least the following Dayhoff sequences, MTY20B11_13 (a mycobacterium tuberculosis peptide), F64471, AE000690_6, XLU16364_1, E43259 (H⁺-transporting ATP synthase) and PIGSLADRXE_1 (MHC class II histocompatibility antigen).

EXAMPLE 58: Isolation of cDNA clones Encoding Human PRO1074

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence (Incyte cluster sequence No. 42586). This cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70

(or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, Univ. of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56251.

In light of an observed sequence homology between the DNA56251 consensus sequence and an EST sequence encompassed within the Merck EST clone no. AA081912, the Merck EST clone AA081912 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 136 and is the full-length DNA sequence for PRO1074. Clone DNA57704-1452 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209953.

The entire nucleotide sequence of DNA57704-1452 is shown in Figure 136 (SEQ ID NO:208). Clone DNA57704-1452 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 322-324 and ending at the stop codon at nucleotide positions 1315-1317 (Figure 136). The predicted polypeptide precursor is 331 amino acids long (Figure 137). The full-length PRO1074 protein shown in Figure 137 has an estimated molecular weight of about 39,512 Daltons and a pI of about 8.03. Analysis of the full-length PRO1074 sequence shown in Figure 137 (SEQ ID NO:209) evidences the presence of the following features: a transmembrane domain at about amino acids 20 to 39; potential N-glycosylation sites at about amino acids 72 to 75, 154 to 157, 198 to 201, 212 to 215, and 326 to 329; a glycosaminoglycan attachment site at about amino acids 239 to 242, and a Ly-6/u-PAR domain at about amino acids 23 to 36.

Analysis of the amino acid sequence of the full-length PRO1074 polypeptide suggests that it possesses significant sequence similarity to beta 1,3-galactosyltransferase, thereby indicating that PRO1074 may be a novel member of the galactosyltransferase family of proteins. Analysis of the amino acid sequence of the full-length PRO1074 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1074 amino acid sequence and the following Dayhoff sequences: AF029792_1, P_R57433, DMU41449_1, AC000348_14, P_R47479, CET09F5_2, CEF14B6_4, CET15D6_5, CEC54C8_4, and CEE03H4_10.

Clone DNA57704-1452 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209953.

EXAMPLE 59: Isolation of cDNA clones Encoding Human PRO1005

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the LIFESEQ® database, Incyte cluster sequence no. 49243. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., *Methods in Enzymology* 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated

DNA56380.

In light of an observed sequence homology between the DNA56380 consensus sequence and an EST sequence encompassed within the Merck EST clone no. AA256657, the Merck EST clone AA256657 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 138 and is herein designated as DNA57708-1411.

The full length clone shown in Figure 138 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 30-32 and ending at the stop codon found at nucleotide positions 585-587 (Figure 138; SEQ ID NO:210). The predicted polypeptide precursor (Figure 139, SEQ ID NO:211) is 185 amino acids long. PRO1005 has a calculated molecular weight of approximately 20,331 daltons and an estimated pI of approximately 5.85. Clone DNA57708-1411 was deposited with the ATCC June 23, 1998, and is assigned ATCC deposit no. 203021.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 139 (SEQ ID NO:211), evidenced some homology between the PRO1005 amino acid sequence and the following Dayhoff sequences: DDU07187_1, DDU87912_1, CELD1007_14, A42239, DDU42597_1, CYAG_DICDI, S50452, MRKC_KLEPN, P-R41998, and XYNA_RUMFL.

EXAMPLE 60: Isolation of cDNA clones Encoding Human PRO1073

An initial DNA sequence referred to herein as DNA55938 and shown in Figure 142 (SEQ ID NO:214) was identified using a yeast screen, in a human SK-Lu-1 adenocarcinoma cell line cDNA library that preferentially represents the 5' ends of the primary cDNA clones. DNA55938 was then compared to ESTs from public databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA), using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)]. The ESTs were clustered and assembled into a consensus DNA sequence using the computer program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained is designated herein as DNA56411.

In light of an observed sequence homology between the DNA56411 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H86027, the Merck EST clone H86027 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 140.

The full length DNA57710-1451 clone shown in Figure 140 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 345-347 and ending at the stop codon found at nucleotide positions 1242-1244 (Figure 140; SEQ ID NO:212). The predicted polypeptide precursor (Figure 141, SEQ ID NO:213) is 299 amino acids long. PRO1073 has a calculated molecular weight of approximately 34,689 daltons and an estimated pI of approximately 11.49. The PRO1073 polypeptide has the following additional features: a signal peptide at about amino acids 1-31, sequence identity to bZIP transcription factor basic domain signature at about amino acids, a potential N-glycosylation site at about amino acids 2-5, and

sequence identity with protamine P1 proteins at about amino acids 158-183.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 141 (SEQ ID NO:213), revealed some sequence identity between the PRO1073 amino acid sequence and the following Dayhoff sequences: MMU37351_1, ATAC00250510T9J22.10, S59043, ENXNUPR_1, B47328, SR55_DROME, S26650, SON_HUMAN, VIT2_CHICK, and XLC4SRPRT_1.

Clone DNA57710-1451 was deposited with the ATCC on July 1, 1998 and is assigned ATCC deposit no. 203048.

EXAMPLE 61: Isolation of cDNA clones Encoding Human PRO1152

A cDNA clone (DNA57711-1501) encoding a native human PRO1152 polypeptide was identified by employing a yeast screen, in a human infant brain cDNA library that preferentially represents the 5' ends of the primary cDNA clones. Specifically, a yeast screen was employed to identify a cDNA designated herein as DNA55807 (SEQ ID NO:217; see Figure 145).

In light of an observed sequence homology between the DNA55807 sequence and an EST sequence encompassed within the Merck EST clone no. R56756, the Merck EST clone R56756 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 143.

The full-length DNA57711-1501 clone shown in Figure 143 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 58-60 and ending at the stop codon at nucleotide positions 1495-1497 (Figure 143). The predicted polypeptide precursor is 479 amino acids long (Figure 144). The full-length PRO1152 protein shown in Figure 144 has an estimated molecular weight of about 53,602 daltons and a pI of about 8.82. Analysis of the full-length PRO1152 sequence shown in Figure 144 (SEQ ID NO:216) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 28, transmembrane domains from about amino acid 133 to about amino acid 155, from about amino acid 168 to about amino acid 187, from about amino acid 229 to about amino acid 247, from about amino acid 264 to about amino acid 285, from about amino acid 309 to about amino acid 330, from about amino acid 371 to about amino acid 390 and from about amino acid 441 to about amino acid 464, potential N-glycosylation sites from about amino acid 34 to about amino acid 37 and from about amino acid 387 to about amino acid 390 and an amino acid sequence block having homology to a respiratory-chain NADH dehydrogenase subunit from about amino acid 243 to about amino acid 287. Clone DNA57711-1501 has been deposited with ATCC on July 1, 1998 and is assigned ATCC deposit no. 203047.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST-2 sequence alignment analysis of the full-length sequence shown in Figure 144 (SEQ ID NO:216), evidenced significant homology between the PRO1152 amino acid sequence and the following Dayhoff sequences: AF052239_1, SYNN9CGA_1, SFCYTB2_1, GEN12507, P_R11769, MTV025_109, C61168, S43171, P_P61689 and P_P61696.

EXAMPLE 62: Isolation of cDNA clones Encoding Human PRO1136

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 109142. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56039.

In light of an observed sequence homology between the DNA56039 consensus sequence and an EST sequence encompassed within the Merck EST clone no. HSC1NF011, the Merck EST clone HSC1NF011 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 146 and is herein designated as DNA57827-1493.

Clone DNA57827-1493) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 216-218 and ending at the stop codon at nucleotide positions 2112-2114 (Figure 146). The predicted polypeptide precursor is 632 amino acids long (Figure 147). The full-length PRO1136 protein shown in Figure 147 has an estimated molecular weight of about 69,643 daltons and a pI of about 8.5. Analysis of the full-length PRO1136 sequence shown in Figure 147 (SEQ ID NO:219) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 15 and potential N-glycosylation sites from about amino acid 108 to about amino acid 111, from about amino acid 157 to about amino acid 160, from about amino acid 289 to about amino acid 292 and from about amino acid 384 to about amino acid 387. Clone DNA57827-1493 has been deposited with ATCC on July 1, 1998 and is assigned ATCC deposit no. 203045.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 147 (SEQ ID NO:219), evidenced significant homology between the PRO1136 amino acid sequence and the following Dayhoff sequences: AF034746_1, AF034745_1, MMAF000168_19, HSMUPP1_1, AF060539_1, SP97_RAT, I38757, MMU93309_1, CEK01A6_4 and HSA224747_1.

EXAMPLE 63: Isolation of cDNA clones Encoding Human PRO813

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence (Incyte EST cluster sequence no. 45501. The Incyte EST cluster sequence no. 45501 sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)).

Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56400.

5 In light of an observed sequence homology between the DNA56400 consensus sequence and an EST sequence encompassed within the Merck EST clone no. T90592, the Merck EST clone T90592 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 148 and is herein designated DNA57834-1339.

10 The full length clone shown in Figure 148 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 109-111 and ending at the stop codon found at nucleotide positions 637-639 (Figure 149; SEQ ID NO:221). The predicted polypeptide precursor is 176 amino acids long, has a calculated molecular weight of approximately 19,616 daltons and an estimated pI of approximately 7.11. Analysis of the full-length PRO813 sequence shown in Figure 149 (SEQ ID NO:221) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 26 and potential N-myristoylation sites from about amino acid 48 to about amino acid 53, from about amino acid 153 to about
15 amino acid 158, from about amino acid 156 to about amino acid 161 and from about amino acid 167 to about amino acid 172. Clone DNA57834-1339 has been deposited with the ATCC on June 9, 1998 and is assigned ATCC deposit no. 209954.

20 Analysis of the amino acid sequence of the full-length PRO813 polypeptide suggests that it possesses sequence similarity to the pulmonary surfactant-associated protein C. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some degree of homology between the PRO813 amino acid sequence and the following Dayhoff sequences, PSPC_MUSVI, P_P92071, G02964, P_R65489, P_P82977, P_R84555, S55542, MUSIGHAJ_1 and PH1158.

EXAMPLE 64: Isolation of cDNA Clones Encoding Human PRO809

25 Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence. The Incyte EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods
30 in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56418.

35 In light of an observed sequence homology between the DNA56418 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H74302, the Merck EST clone H74302 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 150 and is herein designated DNA57836-1338.

The entire nucleotide sequence of DNA57836-1338 is shown in Figure 150 (SEQ ID NO:222). Clone DNA57836-1338 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 63-65 and ending at the stop codon at nucleotide positions 858-860 of SEQ ID NO:222 (Figure 150). The predicted polypeptide precursor is 265 amino acids long (Figure 151). The full-length PRO809 protein shown in Figure 151 has an estimated molecular weight of about 29,061 daltons and a pI of about 9.18. Figure 151 further shows the approximate positions of the signal peptide and N-glycosylation sites. The corresponding nucleotides can be determined by referencing Figure 150. Clone DNA57836-1338 has been deposited with ATCC on June 23, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO809 polypeptide suggests that it possesses some sequence similarity to the heparin sulfate proteoglycan and to endothelial cell adhesion molecule-1. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced sequence identity between the PRO809 amino acid sequence and the following Dayhoff sequences, PGBM_MOUSE, D82082_1 and PW14158.

EXAMPLE 65: Isolation of cDNA Clones Encoding Human PRO791

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence. The Incyte EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56429.

In light of an observed sequence homology between the DNA56429 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 36367, the Merck EST clone 36367 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 152 and is herein designated DNA57838-1337.

The entire nucleotide sequence of DNA57838-1337 is shown in Figure 152 (SEQ ID NO:224). Clone DNA57838-1337 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 9-11 and ending at the stop codon at nucleotide positions 747-749 of SEQ ID NO:224 (Figure 152). The predicted polypeptide precursor is 246 amino acids long (Figure 153). The full-length PRO791 protein shown in Figure 153 has an estimated molecular weight of about 27,368 daltons and a pI of about 7.45. Figure 153 also shows the approximate locations of the signal peptide, the transmembrane domain, N-glycosylation sites and a region conserved in extracellular proteins. The corresponding nucleotides of one embodiment provided herein can be identified by referencing Figure 152. Clone DNA57838-1337 has been deposited with ATCC on June 23, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and

that the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO791 polypeptide suggests that it has sequence similarity with MHC-I antigens, thereby indicating that PRO791 may be related to MHC-I antigens. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some sequence identity between the PRO791 amino acid sequence and the following Dayhoff sequences, AF034346_1, MMQ1K5_1 and HFE_HUMAN.

EXAMPLE 66: Isolation of cDNA clones Encoding Human PRO1004

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence, Incyte cluster sequence No. 73681. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, Univ. of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated as DNA56516.

In light of an observed sequence homology between the DNA56516 consensus sequence and an EST sequence encompassed within the Merck EST clone no. H43837, the Merck EST clone H43837 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 154.

The full length clone shown in Figure 154 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 119-121 and ending at the stop codon at nucleotide positions 464-466 (Figure 154; SEQ ID NO:226). The predicted polypeptide precursor is 115 amino acids long (Figure 155; SEQ ID NO:227). The full-length PRO1004 protein shown in Figure 155 has an estimated molecular weight of about 13,649 daltons and a pI of about 9.58. Analysis of the full-length PRO1004 sequence shown in Figure 155 (SEQ ID NO:227) evidences the presence of the following features: a signal peptide at about amino acids 1-24, a microbodies C-terminal targeting signal at about amino acids 113-115, a potential N-glycosylation site at about amino acids 71-74, and a domain having sequence identity with dihydrofolate reductase proteins at about amino acids 22-48.

Analysis of the amino acid sequence of the full-length PRO1004 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1004 amino acid sequence and the following Dayhoff sequences: CELR02D3_7, LECI_MOUSE, AF006691_3, SSZ97390_1, SSZ97395_1, and SSZ97400_1.

Clone DNA57844-1410 was deposited with the ATCC on June 23, 1998, and is assigned ATCC deposit no. 203010.

EXAMPLE 67: Isolation of cDNA clones Encoding Human PRO1111

An expressed sequence tag (EST) DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST was identified which had homology to insulin-like growth factor binding protein.

RNA for construction of cDNA libraries was isolated from human fetal brain. The cDNA libraries used to isolate the cDNA clones encoding human PRO1111 were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, 253:1278-1280 (1991)) in the unique XhoI and NotI.

The human fetal brain cDNA libraries (prepared as described above), were screened by hybridization with a synthetic oligonucleotide probe based upon the Incyte EST sequence described above:

5'-CCACCACCTGGAGGTCCTGCAGTTGGGCAGGAAGTCCATCCGGCAGATTG-3' (SEQ ID NO:251).

An identified cDNA clone was sequenced in entirety. The entire nucleotide sequence of PRO1111 is shown in Figure 156 (SEQ ID NO:228). Clone DNA58721-1475 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 57-59 and a stop codon at nucleotide positions 2016-2018 (Figure 156; SEQ ID NO:228). The predicted polypeptide precursor is 653 amino acids long (Figure 157). The transmembrane domains are at positions 21-40 (type II) and 528-548. Clone DNA58721-1475 has been deposited with ATCC and is assigned ATCC deposit no. 203110. The full-length PRO1111 protein shown in Figure 157 has an estimated molecular weight of about 72,717 daltons and a pI of about 6.99.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 157 (SEQ ID NO:229), revealed some sequence identity between the PRO1111 amino acid sequence and the following Dayhoff sequences: A58532, D86983_1, RNPLGPV_1, PGS2_HUMAN, AF038127_1, ALS_MOUSE, GPV_HUMAN, PGS2_BOVIN, ALS_PAPPA and I47020.

EXAMPLE 68: Isolation of cDNA clones Encoding Human PRO1344

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA33790. Based on the DNA33790 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1344.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-AGGTTCGTGATGGAGACAACCGCG-3' (SEQ ID NO:232)

reverse PCR primer 5'-TGTCAAGGACGCACTGCCGTCATG-3' (SEQ ID NO:233)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA33790 sequence which had the following nucleotide sequence

hybridization probe

5'-TGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCC-3' (SEQ ID NO:234)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1344 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1344 (designated herein as DNA58723-1588 [Figure 158, SEQ ID NO:230]); and the derived protein sequence for PRO1344.

The entire nucleotide sequence of DNA58723-1588 is shown in Figure 158 (SEQ ID NO:230). Clone DNA58723-1588 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 26-28 and ending at the stop codon at nucleotide positions 2186-2188 (Figure 158). The predicted polypeptide precursor is 720 amino acids long (Figure 159). The full-length PRO1344 protein shown in Figure 159 has an estimated molecular weight of about 80,199 daltons and a pI of about 7.77. Analysis of the full-length PRO1344 sequence shown in Figure 159 (SEQ ID NO:231) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 23, an EGF-like domain cysteine protein signature sequence from about amino acid 260 to about amino acid 271, potential N-glycosylation sites from about amino acid 96 to about amino acid 99, from about amino acid 279 to about amino acid 282, from about amino acid 316 to about amino acid 319, from about amino acid 451 to about amino acid 454 and from about amino acid 614 to about amino acid 617, an amino acid sequence block having homology to serine proteases, trypsin family from about amino acid 489 to about amino acid 505 and a CUB domain protein profile sequence from about amino acid 150 to about amino acid 166. Clone DNA58723-1588 has been deposited with ATCC on August 18, 1998 and is assigned ATCC deposit no. 203133.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 159 (SEQ ID NO:231), evidenced significant homology between the PRO1344 amino acid sequence and the following Dayhoff sequences: S77063_1, CRAR_MOUSE, P_R74775, P_P90070, P_R09217, P_P70475, HSBMP16_1 and U50330_1.

EXAMPLE 69: Isolation of cDNA clones Encoding Human PRO1109

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA52642. The consensus DNA sequence was obtained by extending using repeated cycles of BLAST and phrap a previously obtained consensus sequence as far as possible using the sources of EST sequences discussed above. Based on the DNA52642 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1109.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCTTACCTCAGAGGCCAGAGCAAGC-3' (SEQ ID NO:237)

reverse PCR primer 5'-GAGCTTCATCCGTTCTGCGTTCACC-3' (SEQ ID NO:238)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA52642 sequence which had the following nucleotide sequence

hybridization probe

5'-CAGGAATGTAAAGCTTTACAGAGGGTCGCCATCCTCGTTCCCCACC-3' (SEQ ID NO:239)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1109 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human SK-Lu-1 adenocarcinoma cell tissue (LIB247).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1109 (designated herein as DNA58737-1473 [Figure 160, SEQ ID NO:235]) and the derived protein sequence for PRO1109.

The entire nucleotide sequence of DNA58737-1473 is shown in Figure 160 (SEQ ID NO:235). Clone DNA58737-1473 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 119-120 and ending at the stop codon at nucleotide positions 1151-1153 (Figure 160). The predicted polypeptide precursor is 344 amino acids long (Figure 161). The full-length PRO1109 protein shown in Figure 161 has an estimated molecular weight of about 40,041 daltons and a pI of about 9.34. Analysis of the full-length PRO1109 sequence shown in Figure 161 (SEQ ID NO:236) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 27, potential N-glycosylation sites from about amino acid 4 to about amino acid 7, from about amino acid 220 to about amino acid 223 and from about amino acid 335 to about amino acid 338 and an amino acid sequence block having homology to xylose isomerase proteins from about amino acid 191 to about amino acid 201. Clone DNA58737-1473 has been deposited with ATCC on August 18, 1998 and is assigned ATCC deposit no. 203136.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 161 (SEQ ID NO:236), evidenced significant homology between the PRO1109 amino acid sequence and the following Dayhoff sequences: HSUDPGAL_1, HSUDPB14_1, NALS_BOVIN, HSU10473_1, CEW02B12_11, YNJ4_CAEEL, AE000738_11, CET24D1_1, S48121 and CEGLY9_1.

EXAMPLE 70: Isolation of cDNA clones Encoding Human PRO1383

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA53961. Based on the DNA53961 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1383.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CATTCCTTACCCTGGACCCAGCTCC-3' (SEQ ID NO:242)

reverse PCR primer 5'-GAAAGGCCACAGCACATCTGGCAG-3' (SEQ ID NO:243)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA53961 sequence which had the following nucleotide sequence

hybridization probe

5'-CCACGACCCGAGCAACTTCCTCAAGACCGACTTGTTTCTCTACAGC-3' (SEQ ID NO:244)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1383 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal brain tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1383 (designated herein as DNA58743-1609 [Figure 162, SEQ ID NO: 240]) and the derived protein sequence for PRO1383.

The entire nucleotide sequence of DNA58743-1609 is shown in Figure 162 (SEQ ID NO:240). Clone DNA58743-1609 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 122-124 and ending at the stop codon at nucleotide positions 1391-1393 (Figure 162). The predicted polypeptide precursor is 423 amino acids long (Figure 163). The full-length PRO1383 protein shown in Figure 163 has an estimated molecular weight of about 46,989 daltons and a pI of about 6.77. Analysis of the full-length PRO1383 sequence shown in Figure 163 (SEQ ID NO:241) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 24, a transmembrane domain from about amino acid 339 to about amino acid 362, and potential N-glycosylation sites from about amino acid 34 to about amino acid 37, from about amino acid 58 to about amino acid 61, from about amino acid 142 to about amino acid 145, from about amino acid 197 to about amino acid 200, from about amino acid 300 to about amino acid 303 and from about amino acid 364 to about amino acid 367. Clone DNA58743-1609 has been deposited with ATCC on August 25, 1998 and is assigned ATCC deposit no. 203154.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 163 (SEQ ID NO:241), evidenced significant homology between the PRO1383 amino acid sequence and the following Dayhoff sequences: NMB_HUMAN, QNR_COTJA, P_W38335, P115_CHICK, P_W38164, A45993_1, MMU70209_1, D83704_1 and P_W39176.

EXAMPLE 71: Isolation of cDNA Clones Encoding Human PRO1003

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence designated herein as 43055. This sequence was then compared to a variety of EST databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater

that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated consen01.

In light of an observed sequence homology between the consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2849382, the Incyte EST clone 2849382 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 164.

The entire nucleotide sequence of DNA58846-1409 is shown in Figure 164 (SEQ ID NO:245). Clone DNA58846-1409 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 41-43 and ending at the stop codon at nucleotide positions 293-295 (Figure 164). The predicted polypeptide precursor is 84 amino acids long (Figure 165). The full-length PRO1003 protein shown in Figure 165 has an estimated molecular weight of about 9,408 daltons and a pI of about 9.28. Analysis of the full-length PRO1003 sequence shown in Figure 165 (SEQ ID NO:246) evidences the presence of a signal peptide at amino acids 1 to about 24, and a cAMP- and cGMP-dependent protein kinase phosphorylation site at about amino acids 58 to about 61. Analysis of the amino acid sequence of the full-length PRO1003 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1003 amino acid sequence and the following Dayhoff sequences: AOPCZA363_3, SRTX_ATREN, A48298, MHVJHMS_1, VGL2_CVMJH, DHDHTC2_2, CORT_RAT, TAL6_HUMAN, P_W14123, and DVUFI_2.

EXAMPLE 72: Isolation of cDNA Clones Encoding Human PRO1108

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA53237.

In light of an observed sequence homology between the DNA53237 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2379881, the Incyte EST clone 2379881 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 166 and is herein designated DNA58848-1472.

The entire nucleotide sequence of DNA58848-1472 is shown in Figure 166 (SEQ ID NO:247). Clone DNA58848-1472 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 77-79 and ending at the stop codon at nucleotide positions 1445-1447 (Figure 166). The predicted polypeptide precursor is 456 amino acids long (Figure 167). The full-length PRO1108 protein shown in Figure 167 has an estimated molecular weight of about 52,071 daltons and a pI of about 9.46. Analysis of the full-length PRO1108 sequence shown in Figure 167 (SEQ ID NO:248) evidences the presence of the following: type II transmembrane domains from about amino acid 22 to about amino acid 42, from about amino acid 156 to about amino acid 176, from about amino acid 180 to about amino acid 199 and from about amino acid 369 to about amino acid 388, potential N-glycosylation sites from about amino acid 247 to about amino acid 250, from about amino acid 327 to about amino acid 330, from about amino acid 328 to about amino acid 331 and from about amino acid 362 to about amino acid 365 and an amino acid block having homology to ER lumen protein retaining receptor protein from about amino acid 153 to about amino acid 190. Clone DNA58848-1472 has

been deposited with ATCC on June 9, 1998 and is assigned ATCC deposit no. 209955.

Analysis of the amino acid sequence of the full-length PRO1108 polypeptide suggests that it possesses significant sequence similarity to the LPAAT protein, thereby indicating that PRO1108 may be a novel LPAAT homolog. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO1108 amino acid sequence and the following Dayhoff sequences, AF015811_1, CER07E3_2, YL35_CAEEL, S73863, CEF59F4_4, P_W06422, MMU41736_1, MTV008_39, P_R99248 and Y67_BPT7.

EXAMPLE 73: Isolation of cDNA Clones Encoding Human PRO1137

The extracellular domain (ECD) sequences (including the secretion signal, if any) of from about 950 known secreted proteins from the Swiss-Prot public protein database were used to search expressed sequence tag (EST) databases. The EST databases included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequence. Using this procedure, Incyte EST No. 3459449, also referred to herein as "DNA7108", was identified as an EST having a BLAST score of 70 or greater that did not encode a known protein.

A consensus DNA sequence was assembled relative to the DNA7108 sequence and other ESTs using repeated cycles of BLAST and the program "phrap" (Phil Green, Univ. of Washington, Seattle, WA). The consensus sequence obtained therefrom is referred to herein as DNA53952.

In light of an observed sequence homology between the DNA53952 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3663102, the Incyte EST clone 3663102 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 168.

The entire nucleotide sequence of DNA58849-1494 is shown in Figure 168 (SEQ ID NO:249). Clone DNA58849-1494 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 77-79 and ending at the stop codon at nucleotide positions 797-799 (Figure 168). The predicted polypeptide precursor is 240 amino acids long (Figure 169). The full-length PRO1137 protein shown in Figure 169 has an estimated molecular weight of about 26,064 daltons and a pI of about 8.65. Analysis of the full-length PRO1137 sequence shown in Figure 169 (SEQ ID NO:250) evidences the presence of a signal peptide at about amino acids 1 to 14 and a potential N-glycosylation site at about amino acids 101-105.

Analysis of the amino acid sequence of the full-length PRO1137 polypeptide suggests that it possesses significant sequence similarity to ribosyltransferase thereby indicating that PRO1137 may be a novel member of the ribosyltransferase family of proteins. Analysis of the amino acid sequence of the full-length PRO1137 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1137 amino acid sequence and the following Dayhoff sequences: MMART5_1, NARG_MOUSE, GEN11909, GEN13794, GEN14406, MMRNART62_1, and P_R41876.

EXAMPLE 74: Isolation of cDNA clones Encoding Human PRO1138

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST sequence, Incyte cluster sequence no. 165212. This cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., *Methods in Enzymology* 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated as DNA54224. The assembly included a proprietary Genentech EST designated herein as DNA49140 (Figure 172; SEQ ID NO:254).

In light of an observed sequence homology between the DNA54224 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3836613, the Incyte EST clone 3836613 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 170 and is the full-length DNA sequence for PRO1138. Clone DNA58850-1495 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209956.

The entire nucleotide sequence of DNA58850-1495 is shown in Figure 170 (SEQ ID NO:252). Clone DNA58850-1495 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 38-40 and ending at the stop codon at nucleotide positions 1043-1045 (Figure 170). The predicted polypeptide precursor is 335 amino acids long (Figure 171). The full-length PRO1138 protein shown in Figure 171 has an estimated molecular weight of about 37,421 Daltons and a pI of about 6.36. Analysis of the full-length PRO1138 sequence shown in Figure 171 (SEQ ID NO:253) evidences the presence of the following features: a signal peptide at about amino acid 1 to about amino acid 22; a transmembrane domain at about amino acids 224 to about 250; a leucine zipper pattern at about amino acids 229 to about 250; and potential N-glycosylation sites at about amino acids 98-101, 142-145, 148-151, 172-175, 176-179, 204-207, and 291-295.

Analysis of the amino acid sequence of the full-length PRO1138 polypeptide suggests that it possesses significant sequence similarity to the CD84, thereby indicating that PRO1138 may be a novel member of the Ig superfamily of polypeptides. More particularly, analysis of the amino acid sequence of the full-length PRO1138 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1138 amino acid sequence and the following Dayhoff sequences: HSU82988_1, HUMLY9_1, P_R97631, P_R97628, P_R97629, P_R97630, CD48_RAT, CD2_HUMAN, P_P93996, and HUMBGP_1.

Clone DNA58850-1495 was deposited with ATCC on June 9, 1998, and is assigned ATCC deposit no. 209956.

EXAMPLE 75: Isolation of cDNA clones Encoding Human PRO1054

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST

cluster sequence from the Incyte database, designated 66212. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55722.

In light of an observed sequence homology between the DNA55722 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 319751, the Incyte EST clone 319751 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 173 and is herein designated as DNA58853-1423.

Clone DNA58853-1423 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 46-48 and ending at the stop codon at nucleotide positions 586-588 (Figure 173). The predicted polypeptide precursor is 180 amino acids long (Figure 174). The full-length PRO1054 protein shown in Figure 174 has an estimated molecular weight of about 20,638 daltons and a pI of about 5.0. Analysis of the full-length PRO1054 sequence shown in Figure 174 (SEQ ID NO:256) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 18, a leucine zipper pattern from about amino acid 155 to about amino acid 176 and amino acid sequence blocks having homology to lipocalin proteins from about amino acid 27 to about amino acid 38 and from about amino acid 110 to about amino acid 120. Clone DNA58853-1423 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203016.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 174 (SEQ ID NO:256), evidenced significant homology between the PRO1054 amino acid sequence and the following Dayhoff sequences: MUP1_MOUSE, MUP6_MOUSE, MUP2_MOUSE, MUP8_MOUSE, MUP5_MOUSE, MUP4_MOUSE, S10124, MUPM_MOUSE, MUP_RAT and ECU70823_1.

EXAMPLE 76: Isolation of cDNA clones Encoding Human PRO994

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 157555. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55728.

In light of an observed sequence homology between the DNA55728 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2860366, the Incyte EST clone 2860366 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 175 and is herein designated as DNA58855-1422.

Clone DNA58855-1422 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 31-33 and ending at the stop codon at nucleotide positions 718-720 (Figure 175). The predicted polypeptide precursor is 229 amino acids long (Figure 176). The full-length PRO994 protein shown in Figure 176 has an estimated molecular weight of about 25,109 daltons and a pI of about 6.83. Analysis of the full-length PRO994 sequence shown in Figure 176 (SEQ ID NO:258) evidences the presence of the following: transmembrane domains from about amino acid 10 to about amino acid 31, from about amino acid 50 to about amino acid 72, from about amino acid 87 to about amino acid 110 and from about amino acid 191 to about amino acid 213, potential N-glycosylation sites from about amino acid 80 to about amino acid 83, from about amino acid 132 to about amino acid 135, from about amino acid 148 to about amino acid 151 and from about amino acid 163 to about amino acid 166 and an amino acid block having homology to TNFR/NGFR cysteine-rich region proteins from about amino acid 4 to about amino acid 11. Clone DNA58855-1422 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203018.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 176 (SEQ ID NO:258), evidenced significant homology between the PRO994 amino acid sequence and the following Dayhoff sequences: AF027204_1, TAL6_HUMAN, ILT4_HUMAN, JC6205, MMU57570_1, S40363, ETU56093_1, S42858, P_R66849 and P_R74751.

EXAMPLE 77: Isolation of cDNA clones Encoding Human PRO812

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 170079. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated as DNA55721.

In light of an observed sequence homology between the DNA55721 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 388964, the Incyte EST clone 388964 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 177 and is herein designated as DNA59205-1421.

Clone DNA59205-1421 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 55-57 and ending at the stop codon at nucleotide positions 304-306 (Figure 177).

The predicted polypeptide precursor is 83 amino acids long (Figure 178). The full-length PRO812 protein shown in Figure 178 has an estimated molecular weight of about 9,201 daltons and a pI of about 9.3. Analysis of the full-length PRO812 sequence shown in Figure 178 (SEQ ID NO:260) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 15, a cAMP- and cGMP-dependent protein kinase phosphorylation site from about amino acid 73 to about amino acid 76 and protein kinase C phosphorylation sites from about amino acid 70 to about amino acid 72 and from about amino acid 76 to about amino acid 78. Clone DNA59205-1421 has been deposited with ATCC on June 23, 1998 and is assigned ATCC deposit no. 203009.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 178 (SEQ ID NO:260), evidenced significant homology between the PRO812 amino acid sequence and the following Dayhoff sequences: P_W35802, P_W35803, PSC1_RAT, S68231, GEN13917, PSC2_RAT, CC10_HUMAN, UTER_RABIT, AF008595_1 and A56413.

EXAMPLE 78: Isolation of cDNA clones Encoding Human PRO1069

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST sequence designated herein as 100727. This sequence was then compared to a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, Univ. of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56001.

In light of an observed sequence homology between the DNA56001 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3533881, the Incyte EST clone 3533881 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 179 and is the full-length DNA sequence for PRO1069. Clone DNA59211-1450 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209960.

The entire nucleotide sequence of DNA59211-1450 is shown in Figure 179 (SEQ ID NO:261). Clone DNA59211-1450 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 197-199 and ending at the stop codon at nucleotide positions 464-466. The predicted polypeptide precursor is 89 amino acids long (Figure 180). The full-length PRO1069 protein shown in Figure 180 has an estimated molecular weight of about 9,433 daltons and a pI of about 8.21. Analysis of the full-length PRO1069 sequence shown in Figure 180 (SEQ ID NO:262) evidences the presence of the following features: a signal peptide sequence at amino acid 1 to about 16; a transmembrane domain at about amino acids 36 to about 59; potential N-myristoylation sites at about amino acids 41-46, 45-50, and 84-89; and homology with extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 at about amino acids 54 to about 66.

Analysis of the amino acid sequence of the full-length PRO1069 polypeptide suggests that it possesses significant sequence similarity to CHIF, thereby indicating that PRO1069 may be a member of the CHIF family of polypeptides. More particularly, analysis of the amino acid sequence of the full-length PRO1069 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1069 amino acid sequence and the following Dayhoff sequences: CHIF_RAT, A55571, PLM_HUMAN, A40533, ATNG_BOVIN, RIC_MOUSE, PETD_SYNY3, VTB1_XENLA, A05009, and S75086.

Clone DNA59211-1450 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209960.

EXAMPLE 79: Isolation of cDNA Clones Encoding Human PRO1129

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence designated herein as 98833. The Incyte EST cluster sequence no. 98833 sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56038.

In light of an observed sequence homology between the DNA56038 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 1335241, the Incyte EST clone 1335241 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 181 and is herein designated DNA59213-1487.

The full length clone shown in Figure 181 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 42-44 and ending at the stop codon found at nucleotide positions 1614-1616 (Figure 181; SEQ ID NO:263). The predicted polypeptide precursor is 524 amino acids long, has a calculated molecular weight of approximately 60,310 daltons and an estimated pI of approximately 7.46. Analysis of the full-length PRO1129 sequence shown in Figure 182 (SEQ ID NO:264) evidences the presence of the following: type II transmembrane domains from about amino acid 13 to about amino acid 32 and from about amino acid 77 to about amino acid 102, a cytochrome P-450 cysteine heme-iron ligand signature sequence from about amino acid 461 to about amino acid 470 and potential N-glycosylation sites from about amino acid 112 to about amino acid 115 and from about amino acid 168 to about amino acid 171. Clone DNA59213-1487 has been deposited with the ATCC on June 9, 1998 and is assigned ATCC deposit no. 209959.

Analysis of the amino acid sequence of the full-length PRO1129 polypeptide suggests that it possesses sequence similarity to the cytochrome P-450 family of proteins. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some degree of homology between the PRO1129 amino acid

sequence and the following Dayhoff sequences, AC004523_1, S45702, AF054821_1 and I53015.

EXAMPLE 80: Isolation of cDNA clones Encoding Human PRO1068

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the LIFESEQ® database, designated Incyte cluster no. 141736. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. One or more of the ESTs was derived from a human mast cell line from a patient with mast cell leukemia. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56094.

In light of an observed sequence homology between the DNA56094 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 004974, the Incyte EST clone 004974 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 183 and is herein designated as DNA59214-1449 (SEQ ID NO:265).

The full length clone shown in Figure 183 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 42-44 and ending at the stop codon found at nucleotide positions 414-416 (Figure 183; SEQ ID NO:265). The predicted polypeptide precursor (Figure 184, SEQ ID NO:266) is 124 amino acids long. PRO1068 has a calculated molecular weight of approximately 14,284 daltons and an estimated pI of approximately 8.14. The PRO1068 polypeptide has the following additional features: a signal peptide sequence at about amino acids 1-20, a urotensin II signature sequence at about amino acids 118-123, a cell attachment sequence at about amino acids 64-66, and a potential cAMP- and cGMP-dependent protein kinase phosphorylation site at about amino acids 112-115.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 184 (SEQ ID NO:266), revealed homology between the PRO1068 amino acid sequence and the following Dayhoff sequences: HALBOP_1, MTV043_36, I50498, and P_R78445

Clone DNA59214-1449 was deposited with the ATCC on July 1, 1998 and is assigned ATCC deposit no.203046.

EXAMPLE 81: Isolation of cDNA clones Encoding Human PRO1066

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST cluster sequence designated herein as 79066. The Incyte EST cluster sequence no. 79066 sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST

databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56121.

In light of an observed sequence homology between the DNA56121 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 1515315, the Incyte EST clone 1515315 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 185 and is herein designated DNA59215-1425.

The full length clone shown in Figure 185 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 176-178 and ending at the stop codon found at nucleotide positions 527-529 (Figure 185; SEQ ID NO:267). The predicted polypeptide precursor is 117 amino acids long, has a calculated molecular weight of approximately 12,911 daltons and an estimated pI of approximately 5.46. Analysis of the full-length PRO1066 sequence shown in Figure 186 (SEQ ID NO:268) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 23, a cAMP- and cGMP-dependent protein kinase phosphorylation site from about amino acid 38 to about amino acid 41 and potential N-myristoylation sites from about amino acid 5 to about amino acid 10, from about amino acid 63 to about amino acid 68 and from about amino acid 83 to about amino acid 88. Clone UNQ524 (DNA59215-1425) has been deposited with the ATCC on June 9, 1998 and is assigned ATCC deposit no. 209961.

Analysis of the amino acid sequence of the full-length PRO1066 polypeptide suggests that it does not possess significant sequence similarity to any known human protein. However, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some degree of homology between the PRO1066 amino acid sequence and the following Dayhoff sequences, MOTI_HUMAN, AF025667_1, MTCY19H9_8 and RABIGKCH_1.

EXAMPLE 82: Isolation of cDNA Clones Encoding Human PRO1184

Use of the signal sequence algorithm described in Example 3 on ESTs from an Incyte database allowed identification a candidate sequence designated herein as DNA56375. This sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56375.

In light of an observed sequence homology between the DNA56375 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 1428374, the Incyte EST clone 1428374 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 187.

The full length clone shown in Figure 187 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 106-108 and ending at the stop codon found at nucleotide positions 532-534 (Figure 187; SEQ ID NO:269). The predicted polypeptide precursor is 142 amino acids long, has a calculated molecular weight of approximately 15,690 daltons and an estimated pI of approximately 9.64. Analysis of the full-length PRO1184 sequence shown in Figure 188 (SEQ ID NO:270) evidences the presence of a signal peptide at about amino acids 1-38. Clone DNA59220-1514 has been deposited with the ATCC on June 9, 1998. It is understood that the deposited clone has the actual sequences and that representations are presented herein.

Analysis of the amino acid sequence of the full-length PRO1184 polypeptide suggests that it possesses some sequence identity with a protein called TIM from *Drosophila virilis*, designated "DVTIMS02_1" in the Dayhoff data base, (version 35.45 SwissProt 35). Other Dayhoff database (version 35.45 SwissProt 35) sequences having some degree of sequence identity with PRO1184 include: WIS1_SCHPO, F002186_1, ATAC00239124 and MSAIPRP_1.

EXAMPLE 83: Isolation of cDNA clones Encoding Human PRO1360

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST sequence from an Incyte database, designated DNA10572. This EST sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank, Merck/Wash. U.) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57314.

In light of an observed sequence homology between the DNA57314 consensus sequence and an EST sequence encompassed within the Merck EST clone no. AA406443, the Merck EST clone AA406443 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 189 and is herein designated as DNA59488-1603.

The full length clone shown in Figure 189 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 54-56 and ending at the stop codon found at nucleotide positions 909-911 (Figure 189; SEQ ID NO:271). The predicted polypeptide precursor (Figure 190, SEQ ID NO:272) is 285 amino acids long. PRO1360 has a calculated molecular weight of approximately 31,433 daltons and an estimated pI of approximately 7.32. Clone DNA59488-1603 was deposited with the ATCC on

August 25, 1998 and is assigned ATCC deposit no. 203157.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 190 (SEQ ID NO:272), revealed sequence identity between the PRO1360 amino acid sequence and the following Dayhoff sequences: UN51_CAEEL, YD4B_SCHPO, AF000634_1, GFO_ZYMMO, YE1J_SCHPO, D86566_1, ZMGFO_1, S76976, PPSA_SYNY3, and CEF28B1_4.

EXAMPLE 84: Isolation of cDNA clones Encoding Human PRO1029

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 18763. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57854.

In light of an observed sequence homology between the DNA57854 consensus sequence and an EST sequence encompassed within the Merck EST clone no. T98880, the Merck EST clone T98880 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 191 and is herein designated as DNA59493-1420.

Clone DNA59493-1420 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 39-41 and ending at the stop codon at nucleotide positions 297-299 (Figure 191). The predicted polypeptide precursor is 86 amino acids long (Figure 192). The full-length PRO1029 protein shown in Figure 192 has an estimated molecular weight of about 9,548 daltons and a pI of about 8.52. Analysis of the full-length PRO1029 sequence shown in Figure 192 (SEQ ID NO:274) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 19, an amino acid block having homology to bacterial rhodopsins retinal binding site protein from about amino acid 50 to about amino acid 61, a prenyl group binding site from about amino acid 83 to about amino acid 86 and a potential N-glycosylation site from about amino acid 45 to about amino acid 48. Clone DNA59493-1420 has been deposited with ATCC on July 1, 1998 and is assigned ATCC deposit no. 203050,

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 192 (SEQ ID NO:274), evidenced significant homology between the PRO1029 amino acid sequence and the following Dayhoff sequences: S66088, AF031815_1, MM4A6L_1, PSEIS52a-1, S17699 and P_R63635.

EXAMPLE 85: Isolation of cDNA clones Encoding Human PRO1139

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST

cluster sequence from the Incyte database, designated 4461. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57312.

The DNA57312 consensus sequence included a 172 nucleotides long public EST (T62095, Merck/University of Washington public database). This EST clone, identified herein as a putative protein coding sequence, was purchased from Merck, and sequenced to provide the coding sequence of PRO1139 (Figure 193). As noted before, the deduced amino acid sequence of DNA59497-1496 shows a significant sequence identity with the deduced amino acid sequence of HSOBRGRP_1. The full-length protein (Figure 194) contains a putative signal peptide between amino acid residues 1 and about 28, and three putative transmembrane domains (approximate amino acid residues 33-52, 71-89, 98-120).

EXAMPLE 86: Isolation of cDNA clones Encoding Human PRO1309

An expressed sequence tag (EST) DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST was identified which showed homology to SLIT.

RNA for construction of cDNA libraries was isolated from human fetal brain tissue. The cDNA libraries used to isolate the cDNA clones encoding human PRO1309 were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI.

The cDNA libraries (prepared as described above), were screened by hybridization with a synthetic oligonucleotide probe derived from the above described Incyte EST sequence:

5'-TCCGTGCAGGGGGACGCCTTTCAGAACTGCGCCGAGTTAAGGAAC-3' (SEQ ID NO:279).

A cDNA clone was isolated and sequenced in entirety. The entire nucleotide sequence of DNA59588-1571 is shown in Figure 195 (SEQ ID NO:277). Clone DNA59588-1571 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 720-722 and a stop codon at nucleotide positions 2286-2288 (Figure 195; SEQ ID NO:277). The predicted polypeptide precursor is 522 amino acids long. The signal peptide is approximately at 1-34 and the transmembrane domain is at approximately 428-450 of SEQ ID NO:278. Clone DNA59588-1571 has been deposited with ATCC and is assigned ATCC deposit no. 203106. The full-length PRO1309 protein shown in Figure 196 has an estimated molecular weight of about 58,614 daltons and a pI of about 7.42.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 196 (SEQ ID NO:278), revealed sequence identity between the PRO1309 amino acid sequence and the following Dayhoff sequences: AB007876_1, GPV_MOUSE, ALS_RAT, P_R85889, LUM_CHICK, AB014462_1, PGS1_CANFA, CEM88_7, A58532 and GEN11209.

EXAMPLE 87: Isolation of cDNA Clones Encoding Human PRO1028

Use of the signal sequence algorithm described in Example 3 above allowed identification of a certain EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA59603.

In light of an observed sequence homology between the DNA59603 sequence and an EST sequence contained within Incyte EST clone no. 1497725, the Incyte EST clone no. 1497725 was purchased and the cDNA insert was obtained and sequenced. It was found that the insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 197 and is herein designated as DNA59603-1419.

The entire nucleotide sequence of DNA59603-1419 is shown in Figure 197 (SEQ ID NO:280). Clone DNA59603-1419 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 21-23 and ending at the stop codon at nucleotide positions 612-614 (Figure 197). The predicted polypeptide precursor is 197 amino acids long (Figure 198). The full-length PRO1028 protein shown in Figure 198 has an estimated molecular weight of about 20,832 daltons and a pI of about 8.74. Clone DNA59603-1419 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analyzing the amino acid sequence of SEQ ID NO:281, the putative signal peptide is at about amino acids 1-19 of SEQ ID NO:281. An N-glycosylation site is at about amino acids 35-38 of SEQ ID NO:281. A C-type lectin domain is at about amino acids 108-117 of SEQ ID NO:281, indicating that PRO513 may be related to or be a lectin. The corresponding nucleotides of these amino acid sequences or others can be routinely determined given the sequences provided herein.

EXAMPLE 88: Isolation of cDNA Clones Encoding Human PRO1027

Use of the signal sequence algorithm described in Example 3 above allowed identification of a certain EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing

homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56399.

5 In light of an observed sequence homology between the DNA56399 sequence and an EST sequence contained within Incyte EST clone no. 937605, the Incyte EST clone no. 937605 was purchased and the cDNA insert was obtained and sequenced. It was found that the insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 199 and is herein designated as DNA59605-1418.

10 The entire nucleotide sequence of DNA59605-1418 is shown in Figure 199 (SEQ ID NO:282). Clone DNA59605-1418 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 31-33 and ending at the stop codon at nucleotide positions 262-264 (Figure 199). The predicted polypeptide precursor is 77 amino acids long (Figure 200). The full-length PRO1027 protein shown in Figure 200 has an estimated molecular weight of about 8,772 daltons and a pI of about 9.62. Clone DNA59605-1418 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains
15 the correct sequence, and the sequences provided herein are based on known sequencing techniques.

Analyzing the amino acid sequence of SEQ ID NO:283, the putative signal peptide is at about amino acids 1-33 of SEQ ID NO:283. The type II fibronectin collagen-binding domain begins at about amino acid 30 of SEQ ID NO:283. The corresponding nucleotides for these amino acid sequences and others can be routinely determined given the sequences provided herein. PRO1027 may be involved in tissue formation or
20 repair.

The following Dayhoff designations appear to have some sequence identity with PRO1027: SFT2_YEAST; ATM3E9_2; A69826; YM16_MARPO; E64896; U60193_2; MTLRAJ205_1; MCU60315_70; SPAS_SHIFL; and S54213.

25 EXAMPLE 89: Isolation of cDNA Clones Encoding Human PRO1107

Use of the signal sequence algorithm described in Example 3 above allowed identification of a certain EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing
30 homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56402.

35 In light of an observed sequence homology between the DNA56402 sequence and an EST sequence contained within Incyte EST clone no. 3203694, the Incyte EST clone no. 3203694 was purchased and the cDNA insert was obtained and sequenced. It was found that the insert encoded a full-length protein. The

sequence of this cDNA insert is shown in Figure 201 and is herein designated as DNA59606-1471.

The entire nucleotide sequence of DNA59606-1471 is shown in Figure 201 (SEQ ID NO:284). Clone DNA59606-1471 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 244-246 and ending at the stop codon at nucleotide positions 1675-1677 of SEQ ID NO:284 (Figure 201). The predicted polypeptide precursor is 477 amino acids long (Figure 202). The full-length PRO1107 protein shown in Figure 202 has an estimated molecular weight of about 54,668 daltons and a pI of about 6.33. Clone DNA59606-1471 has been deposited with ATCC on June 9, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO1107 polypeptide suggests that it possesses significant sequence similarity to phosphodiesterase I/nucleotide pyrophosphatase, human insulin receptor tyrosine kinase inhibitor, alkaline phosphodiesterase and autotaxin, thereby indicating that PRO1107 may have at least one or all of the activities of these proteins, and that PRO1107 is a novel phosphodiesterase. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced sequence identity between the PRO1107 amino acid sequence and at least the following Dayhoff sequences: AF005632_1, P_R79148, RNU78787_1, AF060218_4, A57080 and HUMATXT_1.

EXAMPLE 90: Isolation of cDNA clones Encoding Human PRO1140

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST sequence, Incyte cluster sequence No. 135917. This sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, Univ. of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56416.

In light of an observed sequence homology between DNA56416 and an EST sequence contained within Incyte EST clone no. 3345705, Incyte EST clone no. 3345705 was obtained and its insert sequenced. It was found that the insert encoded a full-length protein. The sequence, designated herein as DNA59607-1497, which is shown in Figure 203, is the full-length DNA sequence for PRO1140. Clone DNA59607-1497 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209946.

The entire nucleotide sequence of DNA59607-1497 is shown in Figure 203 (SEQ ID NO:286). Clone DNA59607-1497 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 210-212 and ending at the stop codon at nucleotide positions 975-977 (Figure 203). The predicted polypeptide precursor is 255 amino acids long (Figure 204). The full-length PRO1140 protein shown in Figure 204 has an estimated molecular weight of about 29,405 daltons and a pI of about 7.64. Analysis of the full-length PRO1140 sequence shown in Figure 204 (SEQ ID NO:287) evidences the presence of three

transmembrane domains at about amino acids 101 to 118, 141 to 161 and 172 to 191.

Analysis of the amino acid sequence of the full-length PRO1140 polypeptide using the Dayhoff database (version 35.45 SwissProt 35) evidenced homology between the PRO1140 amino acid sequence and the following Dayhoff sequences: AF023602_1, AF000368_1, CIN3_RAT, AF003373_1, GEN13279, and AF003372_1.

Clone DNA59607-1497 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209946.

EXAMPLE 91: Isolation of cDNA clones Encoding Human PRO1106

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single Incyte EST sequence. This sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, Univ. of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56423.

In light of an observed sequence homology between DNA56423 and an EST sequence contained within Incyte EST clone no. 1711247, Incyte EST clone no. 1711247 was obtained and its insert sequenced. It was found that the insert encoded a full-length protein. The sequence, designated herein as DNA59609-1470, which is shown in Figure 205, is the full-length DNA sequence for PRO1106. Clone DNA59609-1470 was deposited with the ATCC on June 9, 1998, and is assigned ATCC deposit no. 209963.

The entire nucleotide sequence of DNA59609-1470 is shown in Figure 205 (SEQ ID NO:288). Clone DNA59609-1470 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 61-63 and ending at the stop codon at nucleotide positions 1468-1470 of SEQ ID NO:288 (Figure 205). The predicted polypeptide precursor is 469 amino acids long (Figure 206). The full-length PRO1106 protein shown in Figure 206 has an estimated molecular weight of about 52,689 daltons and a pI of about 8.68. It is understood that the skilled artisan can construct the polypeptide or nucleic acid encoding therefor to exclude any one or more of all of these domains. For example, the transmembrane domain region(s) and/or either of the amino terminal or carboxyl end can be excluded. Clone DNA59609-1470 has been deposited with ATCC on June 9, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO1106 polypeptide suggests that it possesses significant sequence similarity to the peroxisomal ca-dependent solute carrier, thereby indicating that PRO1106 may be a novel transporter. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced sequence identity between the PRO1106 amino acid sequence and at least the following Dayhoff sequences, AF004161_1, IG002N01_25, GDC_BOVIN and BT1_MAIZE.

EXAMPLE 92: Isolation of cDNA clones Encoding Human PRO1291

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 120480. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56425.

In light of an observed sequence homology between the DNA56425 sequence and an EST sequence encompassed within the Incyte EST clone no. 2798803, the Incyte EST clone 2798803 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 207 and is herein designated as DNA59610-1556.

Clone DNA59610-1556 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 61-63 and ending at the stop codon at nucleotide positions 907-909 (Figure 207). The predicted polypeptide precursor is 282 amino acids long (Figure 208). The full-length PRO1291 protein shown in Figure 208 has an estimated molecular weight of about 30,878 daltons and a pI of about 5.27. Analysis of the full-length PRO1291 sequence shown in Figure 208 (SEQ ID NO:291) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 28, a transmembrane domain from about amino acid 258 to about amino acid 281 and potential N-glycosylation sites from about amino acid 112 to about amino acid 115, from about amino acid 160 to about amino acid 163, from about amino acid 190 to about amino acid 193, from about amino acid 196 to about amino acid 199, from about amino acid 205 to about amino acid 208, from about amino acid 216 to about amino acid 219 and from about amino acid 220 to about amino acid 223.. Clone DNA59610-1556 has been deposited with ATCC on June 16, 1998 and is assigned ATCC deposit no. 209990.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 208 (SEQ ID NO:291), evidenced significant homology between the PRO1291 amino acid sequence and the following Dayhoff sequences: HSU90552_1, HSU90144_1, AF033107_1, HSB73_1, HSU90142_1, GGCD80_1, P_W34452, MOG_MOUSE, B39371 and P_R71360.

EXAMPLE 93: Isolation of cDNA clones Encoding Human PRO1105

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul

et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56430.

In light of an observed sequence homology between the DNA56430 sequence and an EST sequence encompassed within the Incyte EST clone no. 1853047, the Incyte EST clone 1853047 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 209 and is herein designated as DNA59612-1466.

The entire nucleotide sequence of DNA59612-1466 is shown in Figure 209 (SEQ ID NO:292). Clone DNA59612-1466 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 28-30 and ending at the stop codon at nucleotide positions 568-570 of SEQ ID NO:292 (Figure 209). The predicted polypeptide precursor is 180 amino acids long (Figure 210). The full-length PRO1105 protein shown in Figure 210 has an estimated molecular weight of about 20,040 daltons and a pI of about 8.35. Clone DNA59612-1466 has been deposited with the ATCC on June 9, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analyzing Figure 210, a signal peptide is at about amino acids 1-19 of SEQ ID NO:293 and transmembrane domains are shown at about amino acids 80-99 and 145-162 of SEQ ID NO:293. It is understood that the skilled artisan could form a polypeptide with all of or any combination or individual selection of these regions. It is also understood that the corresponding nucleic acids can be routinely identified and prepared based on the information provided herein.

EXAMPLE 94: Isolation of cDNA clones Encoding Human PRO511

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56434.

In light of an observed sequence homology between the DNA56434 sequence and an EST sequence encompassed within the Incyte EST clone no. 1227491, the Incyte EST clone 1227491 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 211 and is herein designated as DNA59613-1417.

The entire nucleotide sequence of DNA59613-1417 is shown in Figure 211 (SEQ ID NO:294). Clone DNA59613-1417 contains a single open reading frame with an apparent translational initiation site at nucleotide

positions 233-235 and ending at the stop codon at nucleotide positions 944-946 (Figure 211). The predicted polypeptide precursor is 237 amino acids long (Figure 212). The full-length PRO511 protein shown in Figure 212 has an estimated molecular weight of about 25,284 daltons and a pI of about 5.74. Clone DNA59613-1417 has been deposited with the ATCC. Regarding the sequence, it is understood that the deposited clone contains the correct sequence, and the sequences provided herein are based on known sequencing techniques.

5 Analyzing the amino acid sequence of SEQ ID NO:295, the putative signal peptide is at about amino acids 1-25 of SEQ ID NO:295. The N-glycosylation sites are at about amino acids 45-48, 73-76, 107-110, 118-121, 132-135, 172-175, 175-178 and 185-188 of SEQ ID NO:295. An arthropod defensins conserved region is at about amino acids 176-182 of SEQ ID NO:295. A kringle domain begins at about amino acid 128 of SEQ ID NO:295 and a ly-6/u-PAR domain begins at about amino acid 6 of SEQ ID NO:295. The
10 corresponding nucleotides of these amino acid sequences and others can be routinely determined given the sequences provided herein.

The designations appearing in a Dayhoff database with which PRO511 has some sequence identity are as follows: SSC20F10_1; SF041083; P_W26579; S44208; JC2394; PSTA_DICDI; A27020; S59310; RAG1_RABIT; and MUSBALBC1_1.

15 EXAMPLE 95: Isolation of cDNA clones Encoding Human PRO1104

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a
20 proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle,
25 Washington). The consensus sequence obtained therefrom is herein designated DNA56446.

In light of an observed sequence homology between the DNA56446 sequence and an EST sequence encompassed within the Incyte EST clone no. 2837496, the Incyte EST clone 2837496 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 213 and is herein designated as DNA59616-1465.

30 The entire nucleotide sequence of DNA59616-1465 is shown in Figure 213 (SEQ ID NO:296). Clone DNA59616-1465 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 109-111 and ending at the stop codon at nucleotide positions 1132-1134 of SEQ ID NO:296 (Figure 213). The predicted polypeptide precursor is 341 amino acids long (Figure 214). The full-length PRO1104 protein shown in Figure 214 has an estimated molecular weight of about 36,769 daltons and a pI of about 9.03.
35 Clone DNA59616-1465 has been deposited with ATCC on June 16, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analyzing Figure 214, a signal peptide is at about amino acids 1-22 of SEQ ID NO:297. N-myristoylation sites are at about amino acids 41-46, 110-115, 133-138, 167-172 and 179-184 of SEQ ID NO:297.

EXAMPLE 96: Isolation of cDNA clones Encoding Human PRO1100

5 Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul
10 et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington).

15 In light of an observed sequence homology between the obtained consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2305379, the Incyte EST clone 2305379 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 215 and is herein designated as DNA59619-1464.

20 The entire nucleotide sequence of DNA59619-1464 is shown in Figure 215 (SEQ ID NO:298). Clone DNA59619-1464 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 33-35 and ending at the stop codon at nucleotide positions 993-995 of SEQ ID NO:298 (Figure 215). The predicted polypeptide precursor is 320 amino acids long (Figure 216). The full-length PRO1100 protein shown in Figure 216 has an estimated molecular weight of about 36,475 daltons and a pI of about 7.29. Clone DNA59619-1464 has been deposited with ATCC on July 1, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing
25 techniques.

30 Upon analyzing SEQ ID NO:299, the approximate locations of the signal peptide, the transmembrane domains, an N-glycosylation site, an N-myristoylation site, a CUB domain and an amiloride-sensitive sodium channel domain are present. It is believed that PRO1100 may function as a channel. The corresponding nucleic acids for these amino acids and others can be routinely determined given SEQ ID NO:299..

EXAMPLE 97: Isolation of cDNA clones Encoding Human PRO836

35 Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (Lifeseq®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70

(or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained is herein designated DNA56453.

In light of an observed sequence homology between the DNA56453 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2610075, the Incyte EST clone 2610075 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 217 and is herein designated as DNA59620-1463.

The entire nucleotide sequence of DNA59620-1463 is shown in Figure 217 (SEQ ID NO:300). Clone DNA59620-1463 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 65-67 and ending at the stop codon at nucleotide positions 1448-1450 of SEQ ID NO:300 (Figure 217). The predicted polypeptide precursor is 461 amino acids long (Figure 218). The full-length PRO836 protein shown in Figure 218 has an estimated molecular weight of about 52,085 daltons and a pI of about 5.36. Analysis of the full-length PRO836 sequence shown in Figure 218 (SEQ ID NO:301) evidences the presence of the following: a signal peptide, N-glycosylation sites, N-myristoylation sites, a domain conserved in the YJL126w/YLR351c/yhcX family of proteins, and a region having sequence identity with SLS1. Clone DNA59620-1463 has been deposited with ATCC on June 16, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO836 polypeptide suggests that it possesses some sequence similarity to SLS1, thereby indicating that PRO836 may be involved in protein translocation of the ER. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced some homology between the PRO836 amino acid sequence and at least the following Dayhoff sequences, S58132, SPBC3B9_1, S66714, CRU40057_1 and IMA_CAEEL.

EXAMPLE 98: Isolation of cDNA clones Encoding Human PRO1141

Use of the signal sequence algorithm described in Example 3 above allowed identification of an EST cluster sequence from the Incyte database, designated 11873. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56518.

In light of an observed sequence homology between the DNA56518 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2679995, the Incyte EST clone 2679995 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 219 and is herein designated as DNA59625-1498.

Clone DNA59625-1498 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 204-206 and ending at the stop codon at nucleotide positions 945-947 (Figure 219). The predicted polypeptide precursor is 247 amino acids long (Figure 220). The full-length PRO1141 protein shown in Figure 220 has an estimated molecular weight of about 26,840 daltons and a pI of about 8.19. Analysis of the full-length PRO1141 sequence shown in Figure 220 (SEQ ID NO:303) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 19 and transmembrane domains from about amino acid 38 to about amino acid 57, from about amino acid 67 to about amino acid 83, from about amino acid 117 to about amino acid 139 and from about amino acid 153 to about amino acid 170. Clone DNA59625-1498 has been deposited with ATCC on June 16, 1998 and is assigned ATCC deposit no. 209992.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 220 (SEQ ID NO:303), evidenced significant homology between the PRO1141 amino acid sequence and the following Dayhoff sequences: CEVF36H2L_2, PCRB7PRJ_1, AB000506_1, LEU95008_1, MRU87980_15, YIGM_ECOLI, STU65700_1, GHU62778_1, CYST_SYNY3 and AF009567_1.

EXAMPLE 99: Isolation of cDNA clones Encoding Human PRO1132

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein as DNA35934. Based on the DNA35934 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1132.

PCR primers (forward and reverse) were synthesized:

forward PCR primer: 5'-TCCTGTGACCACCCCTCTAACACC-3' (SEQ ID NO:310) and

reverse PCR primer: 5'-CTGGAACATCTGCTGCCCAGATTC-3' (SEQ ID NO:311).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus sequence which had the following nucleotide sequence:

5'-GTCGGATGACAGCAGCAGCCGCATCATCAATGGATCCGACTGCGATATGC-3' (SEQ ID NO:312).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1132 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1132 and the derived protein sequence for PRO1132.

The entire nucleotide sequence of PRO1132 is shown in Figure 225 (SEQ ID NO:308). Clone DNA59767-1489 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 354-356 and a stop codon at nucleotide positions 1233-1235 (Figure 225; SEQ ID NO:308). The predicted polypeptide precursor is 293 amino acids long. The signal peptide is at about amino acids 1-22 and the histidine active site is at about amino acids 104-109 of SEQ ID NO:309. Clone DNA59767-1489 has been

deposited with ATCC (having the actual sequence rather than representations based on sequencing techniques as presented herein) and is assigned ATCC deposit no. 203108. The full-length PRO1132 protein shown in Figure 226 has an estimated molecular weight of about 32,020 daltons and a pI of about 8.7.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 226 (SEQ ID NO:309), revealed sequence identity between the PRO1132 amino acid sequence and the following Dayhoff sequences: SSU76256_1, P_W10694, MMAE000663_6, AF013988_1, U66061_8, MMAE000665_2, MMAE00066415, MMAE00066414, MMAE000665_4 and MMAE00066412.

EXAMPLE 100: Isolation of cDNA clones Encoding Human NL7 (PRO1346)

A single EST sequence (#1398422) was found in the LIFESEQ[®] database as described in Example 1 above. This EST sequence was renamed as DNA45668. Based on the DNA45668 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for NL7.

PCR primers (forward and reverse) were synthesized:

forward PCR primer: 5'-CACACGTCCAACCTCAATGGGCAG-3' (SEQ ID NO:315)

reverse PCR primer: 5'-GACCAGCAGGGCCAAGGACAAGG-3' (SEQ ID NO:316)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA45668 sequence which had the following nucleotide sequence:

hybridization probe:

5'-GTTCTCTGAGATGAAGATCCGGCCGGTCCGGGAGTACCGCTTAG-3'
(SEQ ID NO:317)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the NL7 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from a human fetal kidney library (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for NL7 (designated herein as DNA59776-1600 [Figure 227, SEQ ID NO:313]) and the derived protein sequence for NL7 (PRO1346).

The entire coding sequence of NL7 (PRO1346) is shown in Figure 227 (SEQ ID NO:313). Clone DNA59776-1600 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 1-3 and an apparent stop codon at nucleotide positions 1384-1386. The predicted polypeptide precursor is 461 amino acids long. The protein contains an apparent type II transmembrane domain at amino acid positions from about 31 to about 50; fibrinogen beta and gamma chains C-terminal domain signature starting at about amino acid position 409, and a leucine zipper pattern starting at about amino acid positions 140, 147, 154 and 161, respectively. Clone DNA59776-1600 has been deposited with ATCC and is assigned ATCC deposit no. 203128. The full-length NL7 protein shown in Figure 228 has an estimated molecular weight of about 50,744 daltons and a pI of about 6.38.

Based on a WU-BLAST2 sequence alignment analysis (using the WU-BLAST2 computer program) of the full-length sequence, NL7 shows significant amino acid sequence identity to a human microfibril-associated glycoprotein (1 MFA4_HUMAN); to known TIE-2 ligands and ligand homologues, ficolin, serum lectin and TGF-1 binding protein.

5 EXAMPLE 101: Isolation of cDNA clones Encoding Human PRO1131

A cDNA sequence isolated in the amylase screen described in Example 2 above is herein designated DNA43546 (see Figure 231; SEQ ID NO:320). The DNA43546 sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing
10 homologues. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA45627.

15 Based on the DNA45627 sequence, oligonucleotide probes were generated and used to screen a human library prepared as described in paragraph 1 of Example 2 above. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science 253:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp.

PCR primers (forward and 2 reverse) were synthesized:

20 forward PCR primer 5'-ATGCAGGCCAAGTACAGCAGCAC-3' (SEQ ID NO:321);
reverse PCR primer 1 5'-CATGCTGACGACTTCCTGCAAGC-3' (SEQ ID NO:322); and
reverse PCR primer 1 5'-CCACACAGTCTCTGCTTCTTGGG-3' (SEQ ID NO:323)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA45627 sequence which had the following nucleotide sequence:

25 hybridization probe
5'-ATGCTGGATGATGATGGGGACACCACCATGAGCCTGCATT-3' (SEQ ID NO:324).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1131 gene using the probe oligonucleotide and one of the PCR primers.

30 A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 144-146, and a stop signal at nucleotide positions 984-986 (Figure 229; SEQ ID NO:318). The predicted polypeptide precursor is 280 amino acids long, has a calculated molecular weight of approximately 31,966 daltons and an estimated pI of approximately 6.26. The transmembrane domain sequence is at about 49-74 of SEQ ID NO:319 and the region having sequence identity
35 with LDL receptors is about 50-265 of SEQ ID NO:319. PRO1131 contains potential N-linked glycosylation sites at amino acid positions 95-98 and 169-172 of SEQ ID NO:319. Clone DNA59777-1480 has been deposited with the ATCC and is assigned ATCC deposit no. 203111.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 230 (SEQ ID NO:319), evidenced some sequence identity between the PRO1131 amino acid sequence and the following Dayhoff sequences: AB010710_1, I49053, I49115, RNU56863_1, LY4A_MOUSE, I55686, MMU56404_1, I49361, AF030313_1 and MMU09739_1.

EXAMPLE 102: Isolation of cDNA clones Encoding Human PRO1281

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein as DNA35720. Based on the DNA35720 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1281.

PCR primers (forward and reverse) were synthesized:

forward PCR primers:

5'-TGGAAGGCTGCCGCAACGACAATC-3' (SEQ ID NO:327);

5'-CTGATGTGGCCGATGTTCTG-3' (SEQ ID NO:328); and

5'-ATGGCTCAGTGTGCAGACAG-3' (SEQ ID NO:329).

reverse PCR primers:

5'-GCATGCTGCTCCGTGAAGTAGTCC-3' (SEQ ID NO:330); and

5'-ATGCATGGGAAAGAAGGCCTGCCC-3' (SEQ ID NO:331).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA35720 sequence which had the following nucleotide sequence:

hybridization probe:

5'-TGCACTGGTGACCACGAGGGGGTGCATATAGCCATCTGGAGCTGAG-3' (SEQ ID NO:332).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO1281 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated human fetal liver.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1281 (designated herein as DNA59820-1549 [Figure 232, SEQ ID NO:325]; and the derived protein sequence for PRO1281.

The entire coding sequence of PRO1281 is shown in Figure 232 (SEQ ID NO:325). Clone DNA59820-1549 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 228-230 and an apparent stop codon at nucleotide positions 2553-2555. The predicted polypeptide precursor is 775 amino acids long. The full-length PRO1281 protein shown in Figure 233 has an estimated molecular weight of about 85,481 daltons and a pI of about 6.92. Additional features include a signal peptide at about amino acids 1-15; and potential N-glycosylation sites at about amino acids 138-141 and 361-364.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 233 (SEQ ID NO:326), revealed some sequence

identity between the PRO1281 amino acid sequence and the following Dayhoff sequences: S44860, CET24D1_1, CEC38H2_3, CAC2_HAECO, B3A2_HUMAN, S22373, CEF38A3_2, CEC34F6_2, CEC34F6_3, and CELT22B11_3.

Clone DNA59820-1549 has been deposited with ATCC and is assigned ATCC deposit no. 203129.

5 EXAMPLE 103: Isolation of cDNA clones Encoding Human PRO1064

A cDNA sequence isolated in the amylase screen described in Example 2 above was found, by the WU-BLAST2 sequence alignment computer program, to have no significant sequence identity to any known human protein. This cDNA sequence is herein designated DNA45288. The DNA45288 sequence was then compared to various EST databases including public EST databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify homologous EST sequences. The comparison was performed using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)]. Those comparisons resulting in a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). This consensus sequence is herein designated DNA48609. Oligonucleotide primers based upon the DNA48609 sequence were then synthesized and employed to screen a human fetal kidney cDNA library which resulted in the identification of the DNA59827-1426 clone shown in Figure 234. The cloning vector was pRK5B (pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)), and the cDNA size cut was less than 2800 bp.

20 The oligonucleotide probes employed were as follows:

forward PCR primer 5'-CTGAGACCCTGCAGCACCATCTG-3' (SEQ ID NO:336)

reverse PCR primer 5'-GGTGCTTCTTGAGCCCCACTTAGC-3' (SEQ ID NO:337)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA48609 sequence which had the following nucleotide sequence

25 hybridization probe

5'-AATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCTGT-3' (SEQ ID NO:338)

A full length clone was identified that contained a single open reading frame with an apparent translational initiation site at nucleotide positions 532-534 and a stop signal at nucleotide positions 991-993 (Figure 234, SEQ ID NO:333). The predicted polypeptide precursor is 153 amino acids long, has a calculated molecular weight of approximately 17,317 daltons and an estimated pI of approximately 5.17. Analysis of the full-length PRO1064 sequence shown in Figure 235 (SEQ ID NO:334) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 24, a transmembrane domain from about amino acid 89 to about amino acid 110, an indole-3-glycerol phosphate synthase homology block from about amino acid 74 to about amino acid 105 and a Myb DNA binding domain protein repeat protein homology block from about amino acid 114 to about amino acid 137. Clone DNA59827-1426 has been deposited with ATCC on August 4, 1998 and is assigned ATCC deposit no. 203089.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 235 (SEQ ID NO:334), evidenced homology between the PRO1064 amino acid sequence and the following Dayhoff sequences: MMNP15PRO_1, BP187PLYH_1, CELF42G8_4, MMU58888_1, GEN14270, TUB8_SOLTU, RCN_MOUSE, HUMRBSY79_1, SESENODA_1 and A21467_1.

EXAMPLE 104: Isolation of cDNA clones Encoding Human PRO1379

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein DNA45232. Based on the DNA45232 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1379.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGGACACCGTACCCTGGTATCTGC-3' (SEQ ID NO:341)

reverse PCR primer 5'-CCAACTCTGAGGAGAGCAAGTGGC-3' (SEQ ID NO:342)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA45232 sequence which had the following nucleotide sequence:

hybridization probe

5'-TGTATGTGCACACCCTCACCATCACCTCCAAGGGCAAGGAGAAC-3' (SEQ ID NO:343).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1379 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1379 which is designated herein as DNA59828-1608 and shown in Figure 237 (SEQ ID NO:339); and the derived protein sequence for PRO1379 (SEQ ID NO:340).

The entire coding sequence of PRO1379 is shown in Figure 237 (SEQ ID NO:339). Clone DNA59828-1608 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 10-12 and an apparent stop codon at nucleotide positions 1732-1734. The predicted polypeptide precursor is 574 amino acids long. The full-length PRO1379 protein shown in Figure 238 has an estimated molecular weight of about 65,355 daltons and a pI of about 8.73. Additional features include a signal peptide at about amino acids 1-17 and potential N-glycosylation sites at about amino acids 160-163, 287-290, and 323-326.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 238 (SEQ ID NO:340), revealed some homology between the PRO1379 amino acid sequence and the following Dayhoff sequences: YHY8_YEAST, AF040625_1, HP714394_1, and HIV18U45630_1.

Clone DNA59828-1608 has been deposited with ATCC and is assigned ATCC deposit no. 203158.

EXAMPLE 105: Isolation of cDNA Clones Encoding Human PRO844

An expressed sequence tag (EST) DNA database (LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST was identified which showed sequence identity with aLP. Based on the information and discoveries provided herein, the clone for this EST, Incyte clone no. 2657496 from a cancerous lung library was further examined.

5 DNA sequencing of the insert for this clone gave a sequence (herein designated as DNA59838-1462; SEQ ID NO:344) which includes the full-length DNA sequence for PRO844 and the derived protein sequence for PRO844.

10 The entire nucleotide sequence of DNA59838-1462 is shown in Figure 239 (SEQ ID NO:344). Clone DNA59838-1462 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 5-7 and ending at the stop codon at nucleotide positions 338-340 of SEQ ID NO:344 (Figure 239). The predicted polypeptide precursor is 111 amino acids long (Figure 240). The full-length PRO844 protein shown in Figure 240 has an estimated molecular weight of about 12,050 daltons and a pI of about 5.45. Clone UNQ544 DNA59838-1462 has been deposited with ATCC on June 16, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on
15 known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO844 polypeptide suggests that it possesses significant sequence similarity to serine protease inhibitors, thereby indicating that PRO844 may be a novel proteinase inhibitor. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO844 amino acid sequence and at least the following Dayhoff
20 sequences, ALK1_HUMAN, P_P82403, P_P82402, ELAF_HUMAN and P_P60950.

EXAMPLE 106: Isolation of cDNA Clones Encoding Human PRO848

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety
25 of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a
30 consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55999.

In light of an observed sequence homology between the DNA55999 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2768571, the Incyte EST clone 2768571 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein.
35 The sequence of this cDNA insert is shown in Figure 241 and is herein designated as DNA59839-1461.

The entire nucleotide sequence of DNA59839-1461 is shown in Figure 241 (SEQ ID NO:346). Clone DNA59839-1461 contains a single open reading frame with an apparent translational initiation site at nucleotide

positions 146-148 and ending at the stop codon at nucleotide positions 1946-1948 of SEQ ID NO:346 (Figure 241). The predicted polypeptide precursor is 600 amino acids long (Figure 242). The full-length PRO848 protein shown in Figure 242 has an estimated molecular weight of about 68,536 daltons. Clone DNA59839-1461 has been deposited with ATCC on June 16, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analysis of the amino acid sequence of the full-length PRO848 polypeptide suggests that it may be a novel sialyltransferase. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced sequence identity between the PRO848 amino acid sequence and at least the following Dayhoff sequences, P_R78619 (GalNAc-alpha-2, 6-sialyltransferase), CAAG5_CHICK (alpha-n-acetylgalactosamide alpha-2, 6-sialyltransferase), HSU14550_1, CAG6_HUMAN and P_R63217 (human alpha-2, 3-sialyltransferase).

EXAMPLE 107: Isolation of cDNA Clones Encoding Human PRO1097

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56006.

In light of an observed sequence homology between the DNA56006 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2408105, the Incyte EST clone 2408105 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 243 and is herein designated as DNA59841-1460.

The entire nucleotide sequence of DNA59841-1460 is shown in Figure 243 (SEQ ID NO:348). Clone DNA59841-1460 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 3-5 and ending at the stop codon at nucleotide positions 276-278 of SEQ ID NO:348 (Figure 243). The predicted polypeptide precursor is 91 amino acids long (Figure 244). The full-length PRO1097 protein shown in Figure 244 has an estimated molecular weight of about 10,542 daltons and a pI of about 10.04. Clone DNA59841-1460 has been deposited with ATCC on July 1, 1998. It is understood that the deposited clone has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analyzing Figure 244, the signal peptide is at about amino acids 1-20 of SEQ ID NO:349. The glycoprotease family protein domain starts at about amino acid 56, and the acyltransferase ChoActase/COT/CPT family peptide starts at about amino acid 49 of SEQ ID NO:349.

EXAMPLE 108: Isolation of cDNA clones Encoding Human PRO1153

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56008.

In light of an observed sequence homology between the DNA56008 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2472409, the Incyte EST clone 2472409 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 245 and is herein designated as DNA59842-1502.

The full length clone shown in Figure 245 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 92-94 and ending at the stop codon found at nucleotide positions 683-685 (Figure 245; SEQ ID NO:350). The predicted polypeptide precursor (Figure 246, SEQ ID NO:351) is 197 amino acids long. PRO1153 has a calculated molecular weight of approximately 21,540 daltons and an estimated pI of approximately 8.31. Clone DNA59842-1502 has been deposited with ATCC and is assigned ATCC deposit no. 209982. It is understood that the correct and actual sequence is in the deposited clone while herein are present representations based on current sequencing techniques which may have minor errors.

Based on a WU-BLAST2 sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO1153 shows some amino acid sequence identity to the following Dayhoff designations: S57447; SOYHRGPC_1; S46965; P_P82971; VCPHEROPH_1; EXTN_TOBAC; MLCB2548_9; ANXA_RABIT; JC5437 and SSGP_VOLCA.

EXAMPLE 109: Isolation of cDNA clones Encoding Human PRO1154

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56025.

In light of an observed sequence homology between the DNA56025 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2169375, the Incyte EST clone 2169375 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 247 and is herein designated as DNA59846-1503.

The full length clone shown in Figure 247 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 86-88 and ending at the stop codon found at nucleotide positions 2909-2911 (Figure 247; SEQ ID NO:352). The predicted polypeptide precursor (Figure 248, SEQ ID NO:353) is 941 amino acids long. PRO1154 has a calculated molecular weight of approximately 107,144 daltons and an estimated pI of approximately 6.26. Clone DNA59846-1503 has been deposited with ATCC and is assigned ATCC deposit no. 209978.

Based on a WU-BLAST2 sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO1154 shows sequence identity to at least the following Dayhoff designations: AB011097_1, AMPN_HUMAN, RNU76997_1, 159331, GEN14047, HSU62768_1, P_R51281, CET07F10_1, SSU66371_1, and AMPRE_HUMAN.

EXAMPLE 110: Isolation of cDNA clones Encoding Human PRO1181

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database, designated herein as 82468. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56029.

In light of an observed sequence homology between the DNA56029 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2186536, the Incyte EST clone 2186536 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 249 and is herein designated as DNA59847-1511.

Clone DNA59847-1511 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 17-19 and ending at the stop codon at nucleotide positions 1328-1330 (Figure 249). The predicted polypeptide precursor is 437 amino acids long (Figure 250). The full-length PRO1181 protein shown in Figure 250 has an estimated molecular weight of about 46,363 daltons and a pI of about 6.22. Analysis of the full-length PRO1181 sequence shown in Figure 250 (SEQ ID NO:355) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 15, potential N-glycosylation sites from about amino acid 46 to about amino acid 49, from about amino acid 189 to about amino acid 192 and from about amino acid 382 to about amino acid 385 and amino acid sequence blocks having homology to Ly-6/u-PAR domain proteins from about amino acid 287 to about amino acid 300 and from about amino acid

98 to about amino acid 111. Clone DNA59847-1511 has been deposited with ATCC on August 4, 1998 and is assigned ATCC deposit no. 203098.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 250 (SEQ ID NO:355), evidenced homology between the PRO1181 amino acid sequence and the following Dayhoff sequences: AF041083_1, P_W26579, RNMAGPIAN_1, CELT13C2_2, LMSAP2GN_1, S61882, CEF35C5_12, DP87_DICDI, GIU47631_1 and P_R07092.

EXAMPLE 111: Isolation of cDNA clones Encoding Human PRO1182

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database, designated herein as 146647. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56033.

In light of an observed sequence homology between the DNA56033 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2595195, the Incyte EST clone 2595195 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 251 and is herein designated as DNA59848-1512.

Clone DNA59848-1512 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 67-69 and ending at the stop codon at nucleotide positions 880-882 (Figure 251). The predicted polypeptide precursor is 271 amino acids long (Figure 252). The full-length PRO1182 protein shown in Figure 252 has an estimated molecular weight of about 28,665 daltons and a pI of about 5.33. Analysis of the full-length PRO1182 sequence shown in Figure 252 (SEQ ID NO:357) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 25, an amino acid block having homology to C-type lectin domain proteins from about amino acid 247 to about amino acid 256 and an amino acid sequence block having homology to C1q domain proteins from about amino acid 44 to about amino acid 77. Clone DNA59848-1512 has been deposited with ATCC on August 4, 1998 and is assigned ATCC deposit no. 203088.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 252 (SEQ ID NO:357), evidenced significant homology between the PRO1182 amino acid sequence and the following Dayhoff sequences: PSPD_BOVIN, CL43_BOVIN, CONG_BOVIN, P_W18780, P_R45005, P_R53257 and CELEGAP7_1.

EXAMPLE 112: Isolation of cDNA clones Encoding Human PRO1155

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56102.

In light of an observed sequence homology between the DNA56102 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2858870, the Incyte EST clone 2858870 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 253 and is herein designated as DNA59849-1504.

The full length clone shown in Figure 253 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 158-160 and ending at the stop codon found at nucleotide positions 563-565 (Figure 253; SEQ ID NO:358). The predicted polypeptide precursor (Figure 254, SEQ ID NO:359) is 135 amino acids long. PRO1155 has a calculated molecular weight of approximately 14,833 daltons and an estimated pI of approximately 9.78. Clone DNA59849-1504 has been deposited with ATCC and is assigned ATCC deposit no. 209986. It is understood that the actual clone has the correct sequence whereas herein are only representations which are prone to minor sequencing errors.

Based on a WU-BLAST2 sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO1155 shows some amino acid sequence identity with the following Dayhoff designations: TKNK_BOVIN; PVB19X587_1; AF019049_1; P_W00948; S72864; P_W00949; I62742; AF038501_1; TKNK_HUMAN; and YAT1_RHOBL. Based on the information provided herein, PRO1155 may play a role in providing neuroprotection and cognitive enhancement.

EXAMPLE 113: Isolation of cDNA clones Encoding Human PRO1156

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database, designated herein as 138851. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56261.

In light of an observed sequence homology between the DNA56261 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3675191, the Incyte EST clone 3675191 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 255 and is herein designated as DNA59853-1505.

The full length clone shown in Figure 255 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 212-214 and ending at the stop codon found at nucleotide positions 689-691 (Figure 255; SEQ ID NO:360). The predicted polypeptide precursor (Figure 256, SEQ ID NO:361) is 159 amino acids long. PRO1156 has a calculated molecular weight of approximately 17,476 daltons, an estimated pI of approximately 9.15, a signal peptide sequence at about amino acids 1 to about 22, and potential N-glycosylation sites at about amino acids 27-30 and 41-44.

Clone DNA59853-1505 was deposited with the ATCC on June 16, 1998 and is assigned ATCC deposit no. 209985.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis (using the ALIGN computer program) of the full-length sequence shown in Figure 256 (SEQ ID NO:361), revealed some homology between the PRO1156 amino acid sequence and the following Dayhoff sequences: D45027_1, P_R79914, JC5309, KBF2_HUMAN, AF010144_1, GEN14351, S68681, P_R79915, ZMTAC_3, and HUMCPGO_1.

EXAMPLE 114: Isolation of cDNA Clones Encoding Human PRO1098

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56377.

In light of an observed sequence homology between the DNA56377 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3050917, the Incyte EST clone 3050917 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 257 and is herein designated as DNA59854-1459.

The entire nucleotide sequence of DNA59854-1459 is shown in Figure 257 (SEQ ID NO:362). Clone DNA59854-1459 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 58-60 and ending at the stop codon at nucleotide positions 292-294 of SEQ ID NO:362 (Figure 257). The predicted polypeptide precursor is 78 amino acids long (Figure 258). The full-length PRO1098 protein shown in Figure 258 has an estimated molecular weight of about 8,396 daltons and a pI of about 7.66. Clone DNA59854-1459 has been deposited with ATCC on June 16, 1998. It is understood that the deposited clone

has the actual nucleic acid sequence and that the sequences provided herein are based on known sequencing techniques.

Analyzing Figure 258, a signal peptide appears to be at about amino acids 1-19 of SEQ ID NO:363, an N-glycosylation site appears to be at about amino acids 37-40 of SEQ ID NO:363, and N-myristoylation sites appear to be at about 15-20, 19-24 and 60-65 of SEQ ID NO:363.

EXAMPLE 115: Isolation of cDNA clones Encoding Human PRO1127

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57959.

In light of an observed sequence homology between the DNA57959 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 685126, the Merck EST clone 685126 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 259 and is herein designated as DNA60283-1484.

The full length clone shown in Figure 259 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 126-128 and ending at the stop codon found at nucleotide positions 327-329 (Figure 259; SEQ ID NO:364). The predicted polypeptide precursor (Figure 260, SEQ ID NO:365) is 67 amino acids long including a signal peptide at about 1-29 of SEQ ID NO:365. PRO1127 has a calculated molecular weight of approximately 7,528 daltons and an estimated pI of approximately 4.95. Clone DNA60283-1484 was deposited with the ATCC on July 1, 1998 and is assigned ATCC deposit no. 203043. It is understood that the deposited clone has the actual sequence, whereas representations which may have minor sequencing errors are presented herein.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 260 (SEQ ID NO:365), revealed some homology between the PRO1127 amino acid sequence and the following Dayhoff sequences: AF037218_48, P_W09638, HBA_HETPO, S39821, KR2_EBV, CET20D3_8, HCU37630_1, HS193B12_10, S40012 and TRITUBC_1.

EXAMPLE 116: Isolation of cDNA clones Encoding Human PRO1126

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a

proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56250.

In light of an observed sequence homology between the DNA56250 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 1437250, the Incyte EST clone 1437250 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 261 and is herein designated as DNA60615-1483.

Clone DNA60615-1483 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 110-112 and ending at the stop codon at nucleotide positions 1316-1318 (Figure 261). The predicted polypeptide precursor is 402 amino acids long (Figure 262). The full-length PRO1126 protein shown in Figure 262 has an estimated molecular weight of about 45,921 daltons and a pI of about 8.60. Analysis of the full-length PRO1126 sequence shown in Figure 262 (SEQ ID NO:367) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 25 and potential N-glycosylation sites from about amino acid 66 to about amino acid 69, from about amino acid 138 to about amino acid 141 and from about amino acid 183 to about amino acid 186. Clone DNA60615-1483 has been deposited with ATCC on June 16, 1998 and is assigned ATCC deposit no. 209980.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 262 (SEQ ID NO:367), evidenced significant homology between the PRO1126 amino acid sequence and the following Dayhoff sequences: I73636, NOMR_HUMAN, MMUSMYOC3_1, HS454G6_1, P_R98225, RNU78105_1, RNU72487_1, AF035301_1, CEELC48E7_4 and CEF11C3_3.

EXAMPLE 117: Isolation of cDNA clones Encoding Human PRO1125

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56540.

In light of an observed sequence homology between the DNA56540 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 1486114, the Incyte EST clone 1486114 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein.

The sequence of this cDNA insert is shown in Figure 263 and is herein designated as DNA60615-1483.

The full length clone shown in Figure 263 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 47-49 and ending at the stop codon found at nucleotide positions 1388-1390 (Figure 263; SEQ ID NO:368). The predicted polypeptide precursor (Figure 264, SEQ ID NO:369) is 447 amino acids long. PRO1125 has a calculated molecular weight of approximately 49,798 daltons and an estimated pI of approximately 9.78. Clone DNA60619-1482 has been deposited with ATCC and is assigned ATCC deposit no. 209993. It is understood that the clone has the actual sequence and that the sequences herein are representations based on current techniques which may be prone to minor errors.

Based on a WU-BLAST2 sequence alignment analysis (using the ALIGN computer program) of the full-length sequence, PRO1125 shows some sequence identity with the following Dayhoff designations: RCO1_NEUCR; S58306; PKWA_THECU; S76086; P_R85881; HET1_PODAN; SPU92792_1; APAF_HUMAN; S76414 and S59317.

EXAMPLE 118: Isolation of cDNA clones Encoding Human PRO1186

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56748.

In light of an observed sequence homology between the DNA56748 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3476792, the Incyte EST clone 3476792 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 265 and is herein designated as DNA60621-1516.

The full length clone shown in Figure 265 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 91-93 and ending at the stop codon found at nucleotide positions 406-408 (Figure 265; SEQ ID NO:370). The predicted polypeptide precursor (Figure 266, SEQ ID NO:371) is 105 amino acids long. The signal peptide is at amino acids 1-19 of SEQ ID NO:371. PRO1186 has a calculated molecular weight of approximately 11,715 daltons and an estimated pI of approximately 9.05. Clone DNA60621-1516 was deposited with the ATCC on August 4, 1998 and is assigned ATCC deposit no. 203091.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 266 (SEQ ID NO:371), revealed some sequence identity between the PRO1186 amino acid sequence and the following Dayhoff sequences: VPRA_DENPO, LFE4_CHICK, AF034208_1, AF030433_1, A55035, COL_RABIT, CELB0507_9, S67826_1, S34665 and

CRU73817_1.

EXAMPLE 119: Isolation of cDNA clones Encoding Human PRO1198

An initial DNA sequence referred to herein as DNA52083 was identified using a yeast screen in a human umbilical vein endothelial cell cDNA library that preferentially represents the 5' ends of the primary cDNA clones. DNA52083 was compared to ESTs from public databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA), using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)]. The ESTs were clustered and assembled into a consensus DNA sequence using the computer program "phrap" (Phil Green, University of Washington, Seattle, Washington). One or more of the ESTs was obtained from human breast skin tissue biopsy. This consensus sequence is designated herein as DNA52780.

In light of an observed sequence homology between the DNA52780 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3852910, the Incyte EST clone 3852910 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 267 and is herein designated as DNA60622-1525.

The full length DNA60622-1525 clone shown in Figure 267 (SEQ ID NO:372) contained a single open reading frame with an apparent translational initiation site at nucleotide positions 54 to 56 and ending at the stop codon found at nucleotide positions 741 to 743. The predicted polypeptide precursor, which is shown in Figure 268 (SEQ ID NO:373), is 229 amino acids long. PRO1198 has a calculated molecular weight of approximately 25,764 daltons and an estimated pI of approximately 9.17. There is a signal peptide sequence at about amino acids 1 through 34. There is sequence identity with glycosyl hydrolases family 31 protein at about amino acids 142 to about 175.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 268 (SEQ ID NO:373), revealed some homology between the PRO1198 amino acid sequence and the following Dayhoff sequences: ATF6H11_6, UCRI_RAT, TOBSUP2NT_1, RCUERF3_1, AMU88186_1, P_W22485, S56579, AF040711_1, DPP4_PIG.

Clone DNA60622-1525 was been deposited with the ATCC on August 4, 1998, and is assigned ATCC deposit no. 203090.

EXAMPLE 120: Isolation of cDNA clones Encoding Human PRO1158

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle,

Washington). The consensus sequence obtained therefrom is herein designated DNA57248.

In light of an observed sequence homology between the DNA57248 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2640776, the Incyte EST clone 2640776 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 269 and is herein designated as DNA60625-1507.

5 The full length clone shown in Figure 269 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 163 to 165 and ending at the stop codon found at nucleotide positions 532 to 534 (Figure 269; SEQ ID NO:374). The predicted polypeptide precursor (Figure 270, SEQ ID NO:375) is 123 amino acids long. PRO1158 has a calculated molecular weight of approximately 13,113 daltons and an estimated pI of approximately 8.53. Additional features include a signal peptide sequence at about amino acids 1-19, a transmembrane domain at about amino acids 56-80, and a potential N-glycosylation site at about amino acids 36-39. Clone DNA60625-1507 was deposited with the ATCC on June 16, 1998 and is assigned ATCC deposit no. 209975.

15 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 270 (SEQ ID NO:375), revealed some homology between the PRO1158 amino acid sequence and the following Dayhoff sequences: ATAC00310510F18A8.10, P_R85151, PHS2_SOLTU, RNMHCIBAC_1, RNA1FMHC_1, I68771, RNRT1A10G_1, PTPA_HUMAN, HUMGACA_1, and CHKPTPA_1.

EXAMPLE 121: Isolation of cDNA clones Encoding Human PRO1159

20 Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57221.

30 In light of an observed sequence homology between the DNA57221 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 376776, the Incyte EST clone 376776 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 271 and is herein designated as DNA60627-1508.

35 Clone DNA60627-1508 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 92-94 and ending at the stop codon at nucleotide positions 362-364 (Figure 271). The predicted polypeptide precursor is 90 amino acids long (Figure 272). The full-length PRO1159 protein shown in Figure 272 has an estimated molecular weight of about 9,840 daltons and a pI of about 10.13. Analysis of the full-length PRO1159 sequence shown in Figure 272 (SEQ ID NO:377) evidences the presence

of the following: a signal peptide from about amino acid 1 to about amino acid 15 and a potential N-glycosylation site from about amino acid 38 to about amino acid 41. Clone DNA60627-1508 has been deposited with ATCC on August 4, 1998 and is assigned ATCC deposit no. 203092.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 272 (SEQ ID NO:377), evidenced significant homology between the PRO1159 amino acid sequence and the following Dayhoff sequences: AF016494_6, AF036708_20, DSSCUTE_1, D89100_1, S28060, MEFA_XENLA, AF020798_12, G70065, E64423, JQ2005.

EXAMPLE 122: Isolation of cDNA clones Encoding Human PRO1124

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56035.

In light of an observed sequence homology between the DNA56035 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 2767646, the Incyte EST clone 2767646 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 273 and is herein designated as DNA60629-1481.

The full length clone shown in Figure 273 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 25-27 and ending at the stop codon found at nucleotide positions 2782-2784 (Figure 273; SEQ ID NO:378). The predicted polypeptide precursor (Figure 274, SEQ ID NO:379) is 919 amino acids long. PRO1124 has a calculated molecular weight of approximately 101,282 daltons and an estimated pI of approximately 5.37. Clone DNA60629-1481 has been deposited with the ATCC and is assigned ATCC deposit no. 209979. It is understood that the deposited clone has the actual sequence, whereas only representations based on current sequencing techniques which may include normal and minor errors, are provided herein.

Based on a WU-BLAST2 sequence alignment analysis of the full-length sequence, PRO1124 shows significant amino acid sequence identity to a chloride channel protein and to ECAM-1. Specifically, the following Dayhoff designations were identified as having sequence identity with PRO1124: ECLC_BOVIN, AF001261_1, P_W06548, SSC6A10_1, AF004355_1, S76691, AF017642, BYU06866_2, CSA_DICDI and SAU47139_2.

EXAMPLE 123: Isolation of cDNA clones Encoding Human PRO1287

An expressed sequence tag (EST) DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) was searched and an EST was identified which showed homology to the fringe protein. This EST sequence was then compared to various EST databases including public EST databases (e.g., GenBank), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify homologous EST sequences. The comparison was performed using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)]. Those comparisons resulting in a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). This consensus sequence obtained is herein designated DNA40568.

Based on the DNA40568 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1287. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, *supra*, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CTCGGGGAAAGGGACTTGATGTTGG-3' (SEQ ID NO:382)

reverse PCR primer 1 5'-GCGAAGGTGAGCCTCTATCTCGTGCC-3' (SEQ ID NO:383)

reverse PCR primer 2 5'-CAGCCTACACGTATTGAGG-3' (SEQ ID NO:384)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA40568 sequence which had the following nucleotide sequence

hybridization probe

5'-CAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCC-3' (SEQ ID NO:385).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO1287 gene using the probe oligonucleotide and one of the PCR primers.

RNA for construction of the cDNA libraries was isolated from human bone marrow tissue. The cDNA libraries used to isolated the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1287 (designated herein as DNA61755-1554 [Figure 275, SEQ ID NO:380]) and the derived protein sequence for PRO1287.

The entire nucleotide sequence of DNA61755-1554 is shown in Figure 275 (SEQ ID NO:380). The full length clone contained a single open reading frame with an apparent translational initiation site at nucleotide positions 655-657 and a stop signal at nucleotide positions 2251-2253 (Figure 275, SEQ ID NO:380). The predicted polypeptide precursor is 532 amino acids long, has a calculated molecular weight of approximately 61,351 daltons and an estimated pI of approximately 8.77. Analysis of the full-length PRO1287 sequence shown in Figure 276 (SEQ ID NO:381) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 27 and potential N-glycosylation sites from about amino acid 315 to about amino acid 318 and from about amino acid 324 to about amino acid 327. Clone DNA61755-1554 has been deposited with ATCC on August 11, 1998 and is assigned ATCC deposit no. 203112.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 276 (SEQ ID NO:381), evidenced significant homology between the PRO1287 amino acid sequence and the following Dayhoff sequences: CET24D1_1, EZRI_BOVIN, GGU19889_1, CC3_YEAST, S74244, NALS_MOUSE, MOES_PIG, S28660, S44860 and YNA4_CAEL.

EXAMPLE 124: Isolation of cDNA clones Encoding Human PRO1312

DNA55773 was identified in a human fetal kidney cDNA library using a yeast screen that preferentially represents the 5' ends of the primary cDNA clones. Based on the DNA55773 sequence, oligonucleotides were synthesized for use as probes to isolate a clone of the full-length coding sequence for PRO1312.

The full length DNA61873-1574 clone shown in Figure 277 (SEQ ID NO:386) contained a single open reading frame with an apparent translational initiation site at nucleotide positions 7-9 and ending at the stop codon found at nucleotide positions 643-645. The predicted polypeptide precursor is 212 amino acids long (Figure 278, SEQ ID NO:387). PRO1312 has a calculated molecular weight of approximately 24,024 daltons and an estimated pI of approximately 6.26. Other features include a signal peptide at about amino acids 1-14; a transmembrane domain at about amino acids 141-160, and potential N-glycosylation sites at about amino acids 76-79 and 93-96.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 278 (SEQ ID NO:387), revealed some homology between the PRO1312 amino acid sequence and the following Dayhoff sequences: GCINTALPH_1, GIBMUC1A_1, P_R96298, AF001406_1, PVU88874_1, P_R85151, AF041409_1, CELC50F2_7, C45875, and AB009510_21.

Clone DNA61873-1574 has been deposited with ATCC and is assigned ATCC deposit no. 203132.

EXAMPLE 125: Isolation of cDNA clones Encoding Human PRO1192

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is designated herein DNA35924. Based on the DNA35924 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1192.

PCR primers (forward and reverse) were synthesized:

forward PCR primer: 5'-CCGAGGCCATCTAGAGGCCAGAGC-3' (SEQ ID NO:390)

reverse PCR primer: 5'-ACAGGCAGAGCCAATGGCCAGAGC-3' (SEQ ID NO:391).

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA35924 sequence which had the following nucleotide sequence:

hybridization probe:

5'-GAGAGGACTGCGGGAGTTTGGGACCTTTGTGCAGACGTGCTCATG-3' (SEQ ID NO:392).

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1192 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal liver and spleen tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1192 designated herein as DNA62814-1521 and shown in Figure 279 (SEQ ID NO:388); and the derived protein sequence for PRO1192 which is shown in Figure 280 (SEQ ID NO:389).

The entire coding sequence of PRO1192 is shown in Figure 279 (SEQ ID NO:388). Clone DNA62814-1521 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 121-123 and an apparent stop codon at nucleotide positions 766-768. The predicted polypeptide precursor is 215 amino acids long. The predicted polypeptide precursor has the following features: a signal peptide at about amino acids 1-21; a transmembrane domain at about amino acids 153-176; potential N-glycosylation sites at about amino acids 39-42 and 118-121; and homology with myelin P0 proteins at about amino acids 27-68 and 99-128 of Figure 280. The full-length PRO1192 protein shown in Figure 280 has an estimated molecular weight of about 24,484 daltons and a pI of about 6.98.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 280 (SEQ ID NO:389), revealed homology between the PRO1192 amino acid sequence and the following Dayhoff sequences: GEN12838, MYP0_HUMAN, AF049498_1, GEN14531, P_W14146, HS46KDA_1, CINB_RAT, OX2G_RAT, D87018_1, and D86996_2.

Clone DNA62814-1521 was deposited with the ATCC on August 4, 1998, and is assigned ATCC deposit no. 203093.

EXAMPLE 126: Isolation of cDNA clones Encoding Human PRO1160

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described

in Example 1 above This consensus sequence is herein designated DNA40650. Based on the DNA40650 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1160.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-GCTCCCTGATCTTCATGTCACCACC-3' (SEQ ID NO:395) .
reverse PCR primer 5'-CAGGGACACACTCTACCATTCTGGGAG-3' (SEQ ID NO:396)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA40650 sequence which had the following nucleotide sequence

hybridization probe

5'-CCATCTTTCTGGTCTCTGCCCAGAATCCGACAACAGCTGCTC-3' (SEQ ID NO:397)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1160 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human breast tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1160 (designated herein as DNA62872-1509 [Figure 281, SEQ ID NO: 393]) and the derived protein sequence for PRO1160.

The entire nucleotide sequence of DNA62872-1509 is shown in Figure 281 (SEQ ID NO:393). Clone DNA62872-1509 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 40-42 and ending at the stop codon at nucleotide positions 310-312 (Figure 281). The predicted polypeptide precursor is 90 amino acids long (Figure 282). The full-length PRO1160 protein shown in Figure 282 has an estimated molecular weight of about 9,039 daltons and a pI of about 4.37. Analysis of the full-length PRO1160 sequence shown in Figure 282 (SEQ ID NO:394) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 19 and a protein kinase C phosphorylation site from about amino acid 68 to about amino acid 70. Clone DNA62872-1509 has been deposited with ATCC on August 4, 1998 and is assigned ATCC deposit no. 203100.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 282 (SEQ ID NO:394), evidenced significant homology between the PRO1160 amino acid sequence and the following Dayhoff sequences: B30305, GEN13490, I53641, S53363, HA34_BRELC, SP96_DICDI, S36326, SSU51197_10, MUC1_XENLA, TCU32448_1 and AF000409_1.

EXAMPLE 127: Isolation of cDNA clones Encoding Human PRO1187

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing

homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57726.

5 In light of an observed sequence homology between the DNA57726 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 358563, the Incyte EST clone 358563 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 283 and is herein designated as DNA62876-1517.

10 The full length clone shown in Figure 283 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 121-123 and ending at the stop codon found at nucleotide positions 481-483 (Figure 283; SEQ ID NO:398). The predicted polypeptide precursor (Figure 284, SEQ ID NO:399) is 120 amino acids long. The signal peptide is at about amino acids 1-17 of SEQ ID NO:399. PRO1187 has a calculated molecular weight of approximately 12,925 daltons and an estimated pI of approximately 9.46. Clone DNA62876-1517 was deposited with the ATCC on August 4, 1998 and is assigned
15 ATCC deposit no. 203095. It is understood that the deposited clone contains the actual sequence and that the representations herein may have minor sequencing errors.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 284 (SEQ ID NO:399), revealed some sequence identity (and therefore some relation) between the PRO1187 amino acid sequence and the following Dayhoff
20 sequences: MGNENDOBX_1, CELF41G3_9, AMPG_STRLI, HSBBOVHERL_2, LEEXTEN10_1, AF029958_1 and P_W04957.

EXAMPLE 128: Isolation of cDNA clones Encoding Human PRO1185

25 Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altshul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70
30 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56426.

In light of an observed sequence homology between the DNA56426 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3284411, the Incyte EST clone 3284411 was purchased
35 and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 285 and is herein designated as DNA62881-1515.

The full length DNA62881-1515 clone shown in Figure 285 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 4-6 and ending at the stop codon found at nucleotide positions 598-600 (Figure 285; SEQ ID NO:400). The predicted polypeptide precursor (Figure 286, SEQ ID NO:401) is 198 amino acids long. The signal peptide is at about amino acids 1-21 of SEQ ID NO:401. PRO1185 has a calculated molecular weight of approximately 22,105 daltons and an estimated pI of approximately 7.73. Clone DNA62881-1515 has been deposited with the ATCC and is assigned ATCC deposit no. 203096.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 286 (SEQ ID NO:401), revealed some sequence identity between the PRO1185 amino acid sequence and the following Dayhoff sequences: TUP1_YEAST, AF041382_1, MAOM_SOLTU, SPPBPHU9_1,I41024, EPCPLCFAIL_1, HSPLEC_1, YKL4_CAEEL, A44643, TGU65922_1.

EXAMPLE 129: Isolation of cDNA clones Encoding Human PRO1345

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA47364. Based on the DNA47364 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1345.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-CCTGGTTATCCCCAGGAAGTCCGAC-3' (SEQ ID NO:404)

reverse PCR primer 5'-CTCTTGCTGCTGCGACAGGCCTC-3' (SEQ ID NO:405)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA47364 sequence which had the following nucleotide sequence

hybridization probe

5'-CGCCCTCCAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGAC-3' (SEQ ID NO:406)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1345 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human breast carcinoma tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1345 (designated herein as DNA64852-1589 [Figure 287, SEQ ID NO:402]) and the derived protein sequence for PRO1345.

The entire nucleotide sequence of DNA64852-1589 is shown in Figure 287 (SEQ ID NO:402). Clone DNA64852-1589 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 7-9 or 34-36 and ending at the stop codon at nucleotide positions 625-627 (Figure 287). The predicted polypeptide precursor is 206 amino acids long (Figure 288). The full-length PRO1345 protein shown in Figure 288 has an estimated molecular weight of about 23,190 daltons and a pI of about 9.40. Analysis of

the full-length PRO1345 sequence shown in Figure 288 (SEQ ID NO:403) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 31 or from about amino acid 10 to about amino acid 31 and a C-type lectin domain signature sequence from about amino acid 176 to about amino acid 190. Clone DNA64852-1589 has been deposited with ATCC on August 18, 1998 and is assigned ATCC deposit no. 203127.

5 An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 288 (SEQ ID NO:403), evidenced significant homology between the PRO1345 amino acid sequence and the following Dayhoff sequences: BTU22298_1, TETN_CARSP, TETN_HUMAN, MABA_RAT, S34198, P_W13144, MACMBPA_1, A46274, PSPD_RAT AND P_R32188.

10 EXAMPLE 130: Isolation of cDNA clones Encoding Human PRO1245

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a
15 proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle,
20 Washington). The consensus sequence obtained therefrom is herein designated DNA56019.

In light of an observed sequence homology between the DNA56019 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 1327836, the Incyte EST clone 1327836 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 289 and is herein designated as DNA64884-1527.

25 The full length clone shown in Figure 289 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 79-81 and ending at the stop codon found at nucleotide positions 391-393 (Figure 289; SEQ ID NO:407). The predicted polypeptide precursor (Figure 290, SEQ ID NO:408) is 104 amino acids long, with a signal peptide sequence at about amino acid 1 to about amino acid 18. PRO1245 has a calculated molecular weight of approximately 10,100 daltons and an estimated pI of
30 approximately 8.76.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 290 (SEQ ID NO:408), revealed some homology between the PRO1245 amino acid sequence and the following Dayhoff sequences: SYA_THETH, GEN11167, MTV044_4, AB011151_1, RLAJ2750_3, SNELIPTRA_1, S63624, C28391, A37907, and
35 S14064.

Clone DNA64884-1245 was deposited with the ATCC on August 25, 1998 and is assigned ATCC deposit no. 203155.

EXAMPLE 131: Isolation of cDNA clones Encoding Human PRO1358

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington).

In light of an observed sequence homology between the consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 88718, the Incyte EST clone 88718 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 291 and is herein designated as DNA64890-1612.

The full length clone shown in Figure 291 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 86 through 88 and ending at the stop codon found at nucleotide positions 1418 through 1420 (Figure 291; SEQ ID NO:409). The predicted polypeptide precursor (Figure 292, SEQ ID NO:410) is 444 amino acids long. The signal peptide is at about amino acids 1-18 of SEQ ID NO:410. PRO1358 has a calculated molecular weight of approximately 50,719 daltons and an estimated pI of approximately 8.82. Clone DNA64890-1612 was deposited with the ATCC on August 18, 1998 and is assigned ATCC deposit no. 203131.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 292 (SEQ ID NO:410), revealed sequence identity between the PRO1358 amino acid sequence and the following Dayhoff sequences: P_W07607, AB000545_1, AB000546_1, A1AT_RAT, AB015164_1, P_P50021, COTR_CAVPO, and HAMHPP_1. The variants claimed in this application exclude these sequences.

EXAMPLE 132: Isolation of cDNA clones Encoding Human PRO1195

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA55716.

In light of an observed sequence homology between the DNA55716 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3252980, the Incyte EST clone 3252980 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 293 and is herein designated as DNA65412-1523.

The full length clone shown in Figure 293 contained a single open reading frame with an apparent translational initiation site at nucleotide positions 58-60 and ending at the stop codon found at nucleotide positions 511-513 (Figure 293; SEQ ID NO:411). The predicted polypeptide precursor (Figure 294, SEQ ID NO:412) is 151 amino acids long. The signal sequence is at about amino acids 1-22 of SEQ ID NO:412. PRO1195 has a calculated molecular weight of approximately 17,277 daltons and an estimated pI of approximately 5.33. Clone DNA65412-1523 was deposited with the ATCC on August 4, 1998 and is assigned ATCC deposit no. 203094.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 294 (SEQ ID NO:412), revealed some sequence identity between the PRO1195 amino acid sequence and the following Dayhoff sequences: MMU28486_1, AF044205_1, P_W31186, CELK03C7_1, F69034, EF1A_METVA, AF024540_1, SSU90353_1, MRSP_STAAU and P_R97680.

EXAMPLE 133: Isolation of cDNA clones Encoding Human PRO1270

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57951.

In light of an observed sequence homology between the DNA57951 consensus sequence and an EST sequence encompassed within the Merck EST clone no. 124878, the Merck EST clone 124878 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 295 and is herein designated as DNA66308-1537.

Clone DNA66308-1537 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 103-105 and ending at the stop codon at nucleotide positions 1042-1044 (Figure 295). The predicted polypeptide precursor is 313 amino acids long (Figure 296). The full-length PRO1270 protein shown in Figure 296 has an estimated molecular weight of about 34,978 daltons and a pI of about 5.71. Analysis of the full-length PRO1270 sequence shown in Figure 296 (SEQ ID NO:414) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 16, a potential N-glycosylation site from about amino acid 163 to about amino acid 166 and glycosaminoglycan attachment sites from about

amino acid 74 to about amino acid 77 and from about amino acid 289 to about amino acid 292. Clone DNA66308-1537 has been deposited with ATCC on August 25, 1998 and is assigned ATCC deposit no. 203159.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 296 (SEQ ID NO:414), evidenced significant homology between the PRO1270 amino acid sequence and the following Dayhoff sequences: XLU86699_1, S49589, FIBA_PARPA, FIBB_HUMAN, P_R47189, AF004326_1, DRTENASCN_1, AF004327_1, P_W01411 and FIBG_BOVIN.

EXAMPLE 134: Isolation of cDNA clones Encoding Human PRO1271

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57955.

In light of an observed sequence homology between the DNA57955 consensus sequence and an EST sequence encompassed within the Merck EST clone no. AA625350, the Merck EST clone AA625350 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 297 and is herein designated as DNA66309-1538.

Clone DNA66309-1538 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 94-96 and ending at the stop codon at nucleotide positions 718-720 (Figure 297). The predicted polypeptide precursor is 208 amino acids long (Figure 298). The full-length PRO1271 protein shown in Figure 298 has an estimated molecular weight of about 21,531 daltons and a pI of about 8.99. Analysis of the full-length PRO1271 sequence shown in Figure 298 (SEQ ID NO:416) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 31 and a transmembrane domain from about amino acid 166 to about amino acid 187. Clone DNA66309-1538 has been deposited with ATCC on September 15, 1998 and is assigned ATCC deposit no. 203235.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 298 (SEQ ID NO:416), evidenced significant homology between the PRO1271 amino acid sequence and the following Dayhoff sequences: S57180, S63257, AGA1_YEAST, BPU43599_1, YS8A_CAEEL, S67570, LSU54556_2, S70305, VGLX_HSVEB, and D88733_1.

EXAMPLE 135: Isolation of cDNA clones Encoding Human PRO1375

A Merck/Wash. U. database was searched and a Merck EST was identified. This sequence was then put in a program which aligns it with other sequences from the Swiss-Prot public database, public EST databases (e.g., GenBank, Merck/Wash. U.), and a proprietary EST database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or BLAST2 [Altschul et al., Methods in Enzymology, 266:460-480 (1996)] as a comparison of the extracellular domain (ECD) protein sequences to a 6 frame translation of the EST sequences. Those comparisons resulting in a BLAST score of 70 (or in some cases, 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program “phrap” (Phil Green, University of Washington, Seattle, Washington).

A consensus DNA sequence was assembled relative to other EST sequences using phrap. This consensus sequence is designated herein “DNA67003”.

Based on the DNA67003 consensus sequence, the nucleic acid (SEQ ID NO:417) was identified in a human pancreas library. DNA sequencing of the clone gave the full-length DNA sequence for PRO1375 and the derived protein sequence for PRO1375.

The entire coding sequence of PRO1375 is shown in Figure 299 (SEQ ID NO:417). Clone DNA67004-1614 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 104-106 and an apparent stop codon at nucleotide positions 698-700 of SEQ ID NO:417. The predicted polypeptide precursor is 198 amino acids long. The transmembrane domains are at about amino acids 11-28 (type II) and 103-125 of SEQ ID NO:418. Clone DNA67004-1614 has been deposited with ATCC and is assigned ATCC deposit no. 203115. The full-length PRO1375 protein shown in Figure 300 has an estimated molecular weight of about 22,531 daltons and a pI of about 8.47.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 300 (SEQ ID NO:418), revealed sequence identity between the PRO1375 amino acid sequence and the following Dayhoff sequences: AF026198_5, CELR12C12_5, S73465, Y011_MYCPN, S64538_1, P_P8150, MUVSHPO10_1, VSH_MUMPL and CVU59751_5.

EXAMPLE 136: Isolation of cDNA clones Encoding Human PRO1385

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program “phrap” (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA57952.

In light of an observed sequence homology between the DNA57952 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3129630, the Incyte EST clone 3129630 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 301 and is herein designated as DNA68869-1610.

Clone DNA68869-1610 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 26-28 and ending at the stop codon at nucleotide positions 410-412 (Figure 301). The predicted polypeptide precursor is 128 amino acids long (Figure 302). The full-length PRO1385 protein shown in Figure 302 has an estimated molecular weight of about 13,663 daltons and a pI of about 10.97. Analysis of the full-length PRO1385 sequence shown in Figure 302 (SEQ ID NO:420) evidences the presence of the following: a signal peptide from about amino acid 1 to about amino acid 28, and glycosylaminoglycan attachment sites from about amino acid 82 to about amino acid 85 and from about amino acid 91 to about amino acid 94. Clone DNA68869-1610 has been deposited with ATCC on August 25, 1998 and is assigned ATCC deposit no. 203164.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 302 (SEQ ID NO:420), evidenced low homology between the PRO1385 amino acid sequence and the following Dayhoff sequences: CELT14A8_1, LMNACHRA1_1, HXD9_HUMAN, CHKCMLF_1, HS5PP34_2, DMDRING_1, A37107_1, MMLUNGENE_1, PUM_DROME and DMU25117_1.

EXAMPLE 137: Isolation of cDNA clones Encoding Human PRO1387

Use of the signal sequence algorithm described in Example 3 above allowed identification of a single EST cluster sequence from the Incyte database. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., GenBank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA56259.

In light of an observed sequence homology between the DNA56259 consensus sequence and an EST sequence encompassed within the Incyte EST clone no. 3507924, the Incyte EST clone 3507924 was purchased and the cDNA insert was obtained and sequenced. It was found that this insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 303 and is herein designated as DNA68872-1620.

Clone DNA68872-1620 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 85-87 and ending at the stop codon at nucleotide positions 1267-1269 (Figure 303). The predicted polypeptide precursor is 394 amino acids long (Figure 304). The full-length PRO1387 protein shown in Figure 304 has an estimated molecular weight of about 44,339 daltons and a pI of about 7.10. Analysis of the full-length PRO1387 sequence shown in Figure 304 (SEQ ID NO:422) evidences the presence

of the following: a signal peptide from about amino acid 1 to about amino acid 19, a transmembrane domain from about amino acid 275 to about amino acid 296, potential N-glycosylation sites from about amino acid 76 to about amino acid 79, from about amino acid 231 to about amino acid 234, from about amino acid 302 to about amino acid 305, from about amino acid 307 to about amino acid 310 and from about amino acid 376 to about amino acid 379, and amino acid sequence blocks having homology to myelin p0 protein from about amino acid 210 to about amino acid 239 and from about amino acid 92 to about amino acid 121. Clone DNA68872-1620 has been deposited with ATCC on August 25, 1998 and is assigned ATCC deposit no. 203160.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 304 (SEQ ID NO:422), evidenced significant homology between the PRO1387 amino acid sequence and the following Dayhoff sequences: P_W36955, MYP0_HETFR, HS46KDA_1, AF049498_1, MYOO_HUMAN, AF030454_1, A53268, SHPTCRA_1, P_W14146 and GEN12838.

EXAMPLE 138: Isolation of cDNA clones Encoding Human PRO1384

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA54192. Based on the DNA54192 sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO1384.

PCR primers (forward and reverse) were synthesized:

forward PCR primer 5'-TGCAGCCCCTGTGACACAACTGG-3' (SEQ ID NO:425)

reverse PCR primer 5'-CTGAGATAACCGAGCCATCCTCCCAC-3' (SEQ ID NO:426)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the DNA54192 sequence which had the following nucleotide sequence:

hybridization probe

5'-GGAGATAGCTGCTATGGGTTCTTCAGGCACAACCTTAACATGGGAAG-3' (SEQ ID NO:427)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO1384 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal liver.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO1384 (designated herein as DNA71159-1617 [Figure 305, SEQ ID NO:423]; and the derived protein sequence for PRO1384.

The entire coding sequence of PRO1384 is shown in Figure 305 (SEQ ID NO:423). Clone DNA71159-1617 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 182-184 and an apparent stop codon at nucleotide positions 869-871. The predicted polypeptide precursor is 229 amino acids long. The full-length PRO1384 protein shown in Figure 306 has an estimated molecular weight of about 26,650 daltons and a pI of about 8.76. Additional features include a type II

transmembrane domain at about amino acids 32-57, and potential N-glycosylation sites at about amino acids 68-71, 120-123, and 134-137.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using a WU-BLAST2 sequence alignment analysis of the full-length sequence shown in Figure 306 (SEQ ID NO:424), revealed homology between the PRO1384 amino acid sequence and the following Dayhoff sequences: AF054819_1, HSAJ1687_1, AF009511_1, AB010710_1, GEN13595, HSAJ673_1, GEN13961, AB005900_1, LECH_CHICK, AF021349_1, and NK13_RAT.

Clone DNA71159-1617 has been deposited with ATCC and is assigned ATCC deposit no. 203135.

EXAMPLE 139: Use of PRO as a hybridization probe

The following method describes use of a nucleotide sequence encoding PRO as a hybridization probe.

DNA comprising the coding sequence of full-length or mature PRO as disclosed herein is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO) in human tissue cDNA libraries or human tissue genomic libraries.

Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO can then be identified using standard techniques known in the art.

EXAMPLE 140: Expression of PRO in *E. coli*

This example illustrates preparation of an unglycosylated form of PRO by recombinant expression in *E. coli*.

The DNA sequence encoding PRO is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., supra. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

PRO may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO is initially amplified using selected PCR primers. The primers will contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an expression vector, which is used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D.600 of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g (NH₄)₂SO₄, 0.71 g sodium citrate•2H₂O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM MgSO₄) and grown for approximately 20-30 hours at 30°C with shaking. Samples are removed to verify expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

E. coli paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4°C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged at 40,000 rpm in a Beckman Ultracentrifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The clarified extract is loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein are pooled and stored at 4°C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The refolding solution is stirred gently at 4°C for 12-36 hours. The refolding reaction is quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution is filtered through a 0.22 micron filter and acetonitrile is added to 2-10% final concentration. The refolded protein is chromatographed on a Poros R1/H reversed phase column

using a mobile buffer of 0.1 % TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance are analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein are pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

Fractions containing the desired folded PRO polypeptide are pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins are formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 141: Expression of PRO in mammalian cells

This example illustrates preparation of a potentially glycosylated form of PRO by recombinant expression in mammalian cells.

The vector, pRK5 (see EP 307,247, published March 15, 1989), is employed as the expression vector. Optionally, the PRO DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO DNA using ligation methods such as described in Sambrook et al., *supra*. The resulting vector is called pRK5-PRO.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 μ g pRK5-PRO DNA is mixed with about 1 μ g DNA encoding the VA RNA gene [Thimmappaya et al., *Cell*, 31:543 (1982)] and dissolved in 500 μ l of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl_2 . To this mixture is added, dropwise, 500 μ l of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM NaPO_4 , and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200 μ Ci/ml ^{35}S -cysteine and 200 μ Ci/ml ^{35}S -methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO may be introduced into 293 cells transiently using the dextran sulfate method described by Sompayrac et al., *Proc. Natl. Acad. Sci.*, 12:7575 (1981). 293 cells are grown to

maximal density in a spinner flask and 700 μ g pRK5-PRO DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μ g/ml bovine insulin and 0.1 μ g/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO can be expressed in CHO cells. The pRK5-PRO can be transfected into CHO cells using known reagents such as CaPO₄ or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ³⁵S-methionine. After determining the presence of PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO can then be concentrated and purified by any selected method.

Epitope-tagged PRO may also be expressed in host CHO cells. The PRO may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO can then be concentrated and purified by any selected method, such as by Ni²⁺-chelate affinity chromatography.

PRO may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins are fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas et al., Nucl. Acids Res. 24:9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

Twelve micrograms of the desired plasmid DNA is introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect[®] (Quiagen), Dosper[®] or Fugene[®] (Boehringer Mannheim). The cells are grown as described in Lucas et al., supra. Approximately 3 x 10⁷ cells are frozen

in an ampule for further growth and production as described below.

The ampules containing the plasmid DNA are thawed by placement into water bath and mixed by vortexing. The contents are pipetted into a centrifuge tube containing 10 mLs of media and centrifuged at 1000 rpm for 5 minutes. The supernatant is aspirated and the cells are resuspended in 10 mL of selective media (0.2 μ m filtered PS20 with 5 % 0.2 μ m diafiltered fetal bovine serum). The cells are then aliquoted into a 100 mL spinner containing 90 mL of selective media. After 1-2 days, the cells are transferred into a 250 mL spinner filled with 150 mL selective growth medium and incubated at 37°C. After another 2-3 days, 250 mL, 500 mL and 2000 mL spinners are seeded with 3×10^5 cells/mL. The cell media is exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in U.S. Patent No. 5,122,469, issued June 16, 1992 may actually be used. A 3L production spinner is seeded at 1.2×10^6 cells/mL. On day 0, the cell number pH is determined. On day 1, the spinner is sampled and sparging with filtered air is commenced. On day 2, the spinner is sampled, the temperature shifted to 33°C, and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion) taken. Throughout the production, the pH is adjusted as necessary to keep it at around 7.2. After 10 days, or until the viability dropped below 70%, the cell culture is harvested by centrifugation and filtering through a 0.22 μ m filter. The filtrate was either stored at 4°C or immediately loaded onto columns for purification.

For the poly-His tagged constructs, the proteins are purified using a Ni-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc-containing) constructs are purified from the conditioned media as follows. The conditioned medium is pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275 μ L of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 142: Expression of PRO in Yeast

The following method describes recombinant expression of PRO in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO from the ADH2/GAPDH promoter. DNA encoding PRO and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of PRO. For secretion, DNA encoding

PRO can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native PRO signal peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of PRO.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing PRO may further be purified using selected column chromatography resins.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 143: Expression of PRO in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of PRO in Baculovirus-infected insect cells.

The sequence coding for PRO is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding PRO or the desired portion of the coding sequence of PRO such as the sequence encoding the extracellular domain of a transmembrane protein or the sequence encoding the mature protein if the protein is extracellular is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression are performed as described by O'Reilley et al., Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-his tagged PRO can then be purified, for example, by Ni²⁺-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., Nature, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM MgCl₂; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45 μm filter. A Ni²⁺-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A₂₈₀ with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% glycerol, pH 6.0), which elutes nonspecifically bound

protein. After reaching A_{280} baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or Western blot with Ni^{2+} -NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His₁₀-tagged PRO are pooled and dialyzed against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

EXAMPLE 144: Preparation of Antibodies that Bind PRO

This example illustrates preparation of monoclonal antibodies which can specifically bind PRO.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, supra. Immunogens that may be employed include purified PRO, fusion proteins containing PRO, and cells expressing recombinant PRO on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO antibodies.

After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35 % polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

The hybridoma cells will be screened in an ELISA for reactivity against PRO. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against PRO is within the skill in the art.

The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

EXAMPLE 145: Purification of PRO Polypeptides Using Specific Antibodies

Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the

art of protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide is purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

5 Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotechnology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

10 Such an immunoaffinity column is utilized in the purification of PRO polypeptide by preparing a fraction from cells containing PRO polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble PRO polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

15 A soluble PRO polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRO polypeptide (*e.g.*, high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/PRO polypeptide binding (*e.g.*, a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and PRO polypeptide is collected.

EXAMPLE 146: Drug Screening

25 This invention is particularly useful for screening compounds by using PRO polypeptides or binding fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

30 Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with an PRO polypeptide or fragment thereof and assaying (i) for the presence of a complex between the agent and the PRO polypeptide or fragment, or (ii) for the presence of a complex between the PRO polypeptide or fragment and the cell, by methods well known in the art. In such competitive binding assays, the PRO polypeptide or fragment is typically labeled. After suitable incubation, free PRO polypeptide or fragment is

separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to PRO polypeptide or to interfere with the PRO polypeptide/cell complex.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. As applied to a PRO polypeptide, the peptide test compounds are reacted with PRO polypeptide and washed. Bound PRO polypeptide is detected by methods well known in the art. Purified PRO polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding PRO polypeptide specifically compete with a test compound for binding to PRO polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRO polypeptide.

EXAMPLE 147: Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active polypeptide of interest (*i.e.*, a PRO polypeptide) or of small molecules with which they interact, *e.g.*, agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the PRO polypeptide or which enhance or interfere with the function of the PRO polypeptide *in vivo* (*c.f.*, Hodgson, Bio/Technology, 9: 19-21 (1991)).

In one approach, the three-dimensional structure of the PRO polypeptide, or of an PRO polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of the PRO polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells, Biochemistry, 31:7796-7801 (1992) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda *et al.*, J. Biochem., 113:742-746 (1993).

It is also possible to isolate a target-specific antibody, selected by functional assay, as described above, and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced

peptides. The isolated peptides would then act as the pharmacore.

By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

EXAMPLE 148: Stimulation of Heart Neonatal Hypertrophy (Assay 1)

This assay is designed to measure the ability of PRO polypeptides to stimulate hypertrophy of neonatal heart. PRO polypeptides testing positive in this assay are expected to be useful for the therapeutic treatment of various cardiac insufficiency disorders.

Cardiac myocytes from 1-day old Harlan Sprague Dawley rats were obtained. Cells ($180\ \mu\text{l}$ at $7.5 \times 10^4/\text{ml}$, serum $<0.1\%$, freshly isolated) are added on day 1 to 96-well plates previously coated with DMEM/F12 + 4% FCS. Test samples containing the test PRO polypeptide or growth medium only (negative control) ($20\ \mu\text{l}/\text{well}$) are added directly to the wells on day 1. PGF ($20\ \mu\text{l}/\text{well}$) is then added on day 2 at final concentration of $10^{-6}\ \text{M}$. The cells are then stained on day 4 and visually scored on day 5, wherein cells showing no increase in size as compared to negative controls are scored 0.0, cells showing a small to moderate increase in size as compared to negative controls are scored 1.0 and cells showing a large increase in size as compared to negative controls are scored 2.0. A positive result in the assay is a score of 1.0 or greater.

The following polypeptides tested positive in this assay: PRO1312.

EXAMPLE 149: Stimulation of Endothelial Cell Proliferation (Assay 8)

This assay is designed to determine whether PRO polypeptides of the present invention show the ability to stimulate adrenal cortical capillary endothelial cell (ACE) growth. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of conditions or disorders where angiogenesis would be beneficial including, for example, wound healing, and the like (as would agonists of these PRO polypeptides). Antagonists of the PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of cancerous tumors.

Bovine adrenal cortical capillary endothelial (ACE) cells (from primary culture, maximum of 12-14 passages) were plated in 96-well plates at 500 cells/well per 100 microliter. Assay media included low glucose DMEM, 10% calf serum, 2 mM glutamine, and 1X penicillin/streptomycin/fungizone. Control wells included the following: (1) no ACE cells added; (2) ACE cells alone; (3) ACE cells plus VEGF (5 ng/ml); and (4) ACE cells plus FGF (5ng/ml). The control or test sample, (in 100 microliter volumes), was then added to the wells (at dilutions of 1%, 0.1% and 0.01%, respectively). The cell cultures were incubated for 6-7 days at $37^\circ\text{C}/5\%\ \text{CO}_2$. After the incubation, the media in the wells was aspirated, and the cells were washed 1X with PBS. An acid phosphatase reaction mixture (100 microliter; 0.1M sodium acetate, pH 5.5, 0.1% Triton X-100, 10 mM p-nitrophenyl phosphate) was then added to each well. After a 2 hour incubation at 37°C , the reaction was stopped by addition of 10 microliters 1N NaOH. Optical density (OD) was measured on a microplate reader at 405 nm.

The activity of a PRO polypeptide was calculated as the fold increase in proliferation (as determined by the acid phosphatase activity, OD 405 nm) relative to (1) cell only background, and (2) relative to maximum stimulation by VEGF. VEGF (at 3-10 ng/ml) and FGF (at 1-5 ng/ml) were employed as an activity reference for maximum stimulation. Results of the assay were considered "positive" if the observed stimulation was \geq 50% increase over background. VEGF (5 ng/ml) control at 1% dilution gave 1.24 fold stimulation; FGF (5 ng/ml) control at 1% dilution gave 1.46 fold stimulation.

The following PRO polypeptides tested positive in this assay: PRO1154 and PRO1186.

EXAMPLE 150: Inhibition of Vascular Endothelial Growth Factor (VEGF) Stimulated Proliferation of Endothelial Cell Growth (Assay 9)

The ability of various PRO polypeptides to inhibit VEGF stimulated proliferation of endothelial cells was tested. Polypeptides testing positive in this assay are useful for inhibiting endothelial cell growth in mammals where such an effect would be beneficial, e.g., for inhibiting tumor growth.

Specifically, bovine adrenal cortical capillary endothelial cells (ACE) (from primary culture, maximum of 12-14 passages) were plated in 96-well plates at 500 cells/well per 100 microliter. Assay media included low glucose DMEM, 10% calf serum, 2 mM glutamine, and 1X penicillin/streptomycin/fungizone. Control wells included the following: (1) no ACE cells added; (2) ACE cells alone; (3) ACE cells plus 5 ng/ml FGF; (4) ACE cells plus 3 ng/ml VEGF; (5) ACE cells plus 3 ng/ml VEGF plus 1 ng/ml TGF-beta; and (6) ACE cells plus 3 ng/ml VEGF plus 5 ng/ml LIF. The test samples, poly-his tagged PRO polypeptides (in 100 microliter volumes), were then added to the wells (at dilutions of 1%, 0.1% and 0.01%, respectively). The cell cultures were incubated for 6-7 days at 37°C/5% CO₂. After the incubation, the media in the wells was aspirated, and the cells were washed 1X with PBS. An acid phosphatase reaction mixture (100 microliter; 0.1M sodium acetate, pH 5.5, 0.1% Triton X-100, 10 mM p-nitrophenyl phosphate) was then added to each well. After a 2 hour incubation at 37°C, the reaction was stopped by addition of 10 microliters 1N NaOH. Optical density (OD) was measured on a microplate reader at 405 nm.

The activity of PRO polypeptides was calculated as the percent inhibition of VEGF (3 ng/ml) stimulated proliferation (as determined by measuring acid phosphatase activity at OD 405 nm) relative to the cells without stimulation. TGF-beta was employed as an activity reference at 1 ng/ml, since TGF-beta blocks 70-90% of VEGF-stimulated ACE cell proliferation. The results are indicative of the utility of the PRO polypeptides in cancer therapy and specifically in inhibiting tumor angiogenesis. Numerical values (relative inhibition) are determined by calculating the percent inhibition of VEGF stimulated proliferation by the PRO polypeptides relative to cells without stimulation and then dividing that percentage into the percent inhibition obtained by TGF- β at 1 ng/ml which is known to block 70-90% of VEGF stimulated cell proliferation. The results are considered positive if the PRO polypeptide exhibits 30% or greater inhibition of VEGF stimulation of endothelial cell growth (relative inhibition 30% or greater).

The following polypeptide tested positive in this assay: PRO812.

EXAMPLE 151: Stimulatory Activity in Mixed Lymphocyte Reaction (MLR) Assay (Assay 24)

This example shows that certain polypeptides of the invention are active as a stimulator of the proliferation of stimulated T-lymphocytes. Compounds which stimulate proliferation of lymphocytes are useful therapeutically where enhancement of an immune response is beneficial. A therapeutic agent may take the form of antagonists of the polypeptide of the invention, for example, murine-human chimeric, humanized or human antibodies against the polypeptide.

The basic protocol for this assay is described in Current Protocols in Immunology, unit 3.12; edited by J E Coligan, A M Kruisbeek, D H Marglies, E M Shevach, W Strober, National Institutes of Health, Published by John Wiley & Sons, Inc.

More specifically, in one assay variant, peripheral blood mononuclear cells (PBMC) are isolated from mammalian individuals, for example a human volunteer, by leukopheresis (one donor will supply stimulator PBMCs, the other donor will supply responder PBMCs). If desired, the cells are frozen in fetal bovine serum and DMSO after isolation. Frozen cells may be thawed overnight in assay media (37°C, 5% CO₂) and then washed and resuspended to 3x10⁶ cells/ml of assay media (RPMI; 10% fetal bovine serum, 1% penicillin/streptomycin, 1% glutamine, 1% HEPES, 1% non-essential amino acids, 1% pyruvate). The stimulator PBMCs are prepared by irradiating the cells (about 3000 Rads).

The assay is prepared by plating in triplicate wells a mixture of:

100:1 of test sample diluted to 1% or to 0.1%,

50 :1 of irradiated stimulator cells, and

50 :1 of responder PBMC cells.

100 microliters of cell culture media or 100 microliter of CD4-IgG is used as the control. The wells are then incubated at 37°C, 5% CO₂ for 4 days. On day 5, each well is pulsed with tritiated thymidine (1.0 mCi/well; Amersham). After 6 hours the cells are washed 3 times and then the uptake of the label is evaluated.

In another variant of this assay, PBMCs are isolated from the spleens of Balb/c mice and C57B6 mice. The cells are teased from freshly harvested spleens in assay media (RPMI; 10% fetal bovine serum, 1% penicillin/streptomycin, 1% glutamine, 1% HEPES, 1% non-essential amino acids, 1% pyruvate) and the PBMCs are isolated by overlaying these cells over Lympholyte M (Organon Teknika), centrifuging at 2000 rpm for 20 minutes, collecting and washing the mononuclear cell layer in assay media and resuspending the cells to 1x10⁷ cells/ml of assay media. The assay is then conducted as described above.

Positive increases over control are considered positive with increases of greater than or equal to 180% being preferred. However, any value greater than control indicates a stimulatory effect for the test protein.

The following PRO polypeptides tested positive in this assay: PRO826, PRO1068, PRO1184, PRO1346 and PRO1375.

EXAMPLE 152: Retinal Neuron Survival (Assay 52)

This example demonstrates that certain PRO polypeptides have efficacy in enhancing the survival of retinal neuron cells and, therefore, are useful for the therapeutic treatment of retinal disorders or injuries including, for example, treating sight loss in mammals due to retinitis pigmentosum, AMD, etc.

Sprague Dawley rat pups at postnatal day 7 (mixed population: glia and retinal neuronal types) are killed by decapitation following CO₂ anesthesia and the eyes are removed under sterile conditions. The neural retina is dissected away from the pigment epithelium and other ocular tissue and then dissociated into a single cell suspension using 0.25 % trypsin in Ca²⁺, Mg²⁺-free PBS. The retinas are incubated at 37°C for 7-10 minutes after which the trypsin is inactivated by adding 1 ml soybean trypsin inhibitor. The cells are plated at 100,000 cells per well in 96 well plates in DMEM/F12 supplemented with N2 and with or without the specific test PRO polypeptide. Cells for all experiments are grown at 37°C in a water saturated atmosphere of 5% CO₂. After 2-3 days in culture, cells are stained with calcein AM then fixed using 4% paraformaldehyde and stained with DAPI for determination of total cell count. The total cells (fluorescent) are quantified at 20X objective magnification using CCD camera and NIH image software for MacIntosh. Fields in the well are chosen at random.

The effect of various concentration of PRO polypeptides are reported herein where percent survival is calculated by dividing the total number of calcein AM positive cells at 2-3 days in culture by the total number of DAPI-labeled cells at 2-3 days in culture. Anything above 30% survival is considered positive.

The following PRO polypeptides tested positive in this assay using polypeptide concentrations within the range of 0.01 % to 1.0% in the assay: PRO828, PRO826, PRO1068 and PRO1132.

EXAMPLE 153: Rod Photoreceptor Cell Survival (Assay 56)

This assay shows that certain polypeptides of the invention act to enhance the survival/proliferation of rod photoreceptor cells and, therefore, are useful for the therapeutic treatment of retinal disorders or injuries including, for example, treating sight loss in mammals due to retinitis pigmentosum, AMD, etc.

Sprague Dawley rat pups at 7 day postnatal (mixed population: glia and retinal neuronal cell types) are killed by decapitation following CO₂ anesthesia and the eyes are removed under sterile conditions. The neural retina is dissected away from the pigment epithelium and other ocular tissue and then dissociated into a single cell suspension using 0.25% trypsin in Ca²⁺, Mg²⁺-free PBS. The retinas are incubated at 37°C for 7-10 minutes after which the trypsin is inactivated by adding 1 ml soybean trypsin inhibitor. The cells are plated at 100,000 cells per well in 96 well plates in DMEM/F12 supplemented with N₂. Cells for all experiments are grown at 37°C in a water saturated atmosphere of 5% CO₂. After 2-3 days in culture, cells are fixed using 4% paraformaldehyde, and then stained using CellTracker Green CMFDA. Rho 4D2 (ascites or IgG 1:100), a monoclonal antibody directed towards the visual pigment rhodopsin is used to detect rod photoreceptor cells by indirect immunofluorescence. The results are calculated as % survival: total number of calcein - rhodopsin positive cells at 2-3 days in culture, divided by the total number of rhodopsin positive cells at time 2-3 days in culture. The total cells (fluorescent) are quantified at 20x objective magnification using a CCD camera and NIH image software for MacIntosh. Fields in the well are chosen at random.

The following polypeptides tested positive in this assay: PRO536, PRO943, PRO828, PRO826, PRO1068 and PRO1132.

EXAMPLE 154: Induction of c-fos in Endothelial Cells (Assay 34)

This assay is designed to determine whether PRO polypeptides show the ability to induce c-fos in endothelial cells. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of conditions or disorders where angiogenesis would be beneficial including, for example, wound healing, and the like (as would agonists of these PRO polypeptides). Antagonists of the PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of cancerous tumors.

Human venous umbilical vein endothelial cells (HUVEC, Cell Systems) in growth media (50% Ham's F12 w/o GHT: low glucose, and 50% DMEM without glycine: with NaHCO₃, 1% glutamine, 10 mM HEPES, 10% FBS, 10 ng/ml bFGF) were plated on 96-well microtiter plates at a cell density of 1×10^4 cells/well. The day after plating, the cells were starved by removing the growth media and treating the cells with 100 μ l/well test samples and controls (positive control = growth media; negative control = Protein 32 buffer = 10 mM HEPES, 140 mM NaCl, 4% (w/v) mannitol, pH 6.8). The cells were incubated for 30 minutes at 37°C, in 5% CO₂. The samples were removed, and the first part of the bDNA kit protocol (Chiron Diagnostics, cat. #6005-037) was followed, where each capitalized reagent/buffer listed below was available from the kit.

Briefly, the amounts of the TM Lysis Buffer and Probes needed for the tests were calculated based on information provided by the manufacturer. The appropriate amounts of thawed Probes were added to the TM Lysis Buffer. The Capture Hybridization Buffer was warmed to room temperature. The bDNA strips were set up in the metal strip holders, and 100 μ l of Capture Hybridization Buffer was added to each b-DNA well needed, followed by incubation for at least 30 minutes. The test plates with the cells were removed from the incubator, and the media was gently removed using the vacuum manifold. 100 μ l of Lysis Hybridization Buffer with Probes were quickly pipetted into each well of the microtiter plates. The plates were then incubated at 55°C for 15 minutes. Upon removal from the incubator, the plates were placed on the vortex mixer with the microtiter adapter head and vortexed on the #2 setting for one minute. 80 μ l of the lysate was removed and added to the bDNA wells containing the Capture Hybridization Buffer, and pipetted up and down to mix. The plates were incubated at 53°C for at least 16 hours.

On the next day, the second part of the bDNA kit protocol was followed. Specifically, the plates were removed from the incubator and placed on the bench to cool for 10 minutes. The volumes of additions needed were calculated based upon information provided by the manufacturer. An Amplifier Working Solution was prepared by making a 1:100 dilution of the Amplifier Concentrate (20 fm/ μ l) in AL Hybridization Buffer. The hybridization mixture was removed from the plates and washed twice with Wash A. 50 μ l of Amplifier Working Solution was added to each well and the wells were incubated at 53°C for 30 minutes. The plates were then removed from the incubator and allowed to cool for 10 minutes. The Label Probe Working Solution was prepared by making a 1:100 dilution of Label Concentrate (40 pmoles/ μ l) in AL Hybridization Buffer. After the 10-minute cool-down period, the amplifier hybridization mixture was removed and the plates were washed twice with Wash A. 50 μ l of Label Probe Working Solution was added to each well and the wells were incubated at 53°C for 15 minutes. After cooling for 10 minutes, the Substrate was warmed to room

temperature. Upon addition of 3 μ l of Substrate Enhancer to each ml of Substrate needed for the assay, the plates were allowed to cool for 10 minutes, the label hybridization mixture was removed, and the plates were washed twice with Wash A and three times with Wash D. 50 μ l of the Substrate Solution with Enhancer was added to each well. The plates were incubated for 30 minutes at 37°C and RLU was read in an appropriate luminometer.

5 The replicates were averaged and the coefficient of variation was determined. The measure of activity of the fold increase over the negative control (Protein 32/HEPES buffer described above) value was indicated by chemiluminescence units (RLU). The results are considered positive if the PRO polypeptide exhibits at least a two-fold value over the negative buffer control. Negative control = 1.00 RLU at 1.00% dilution. Positive control = 8.39 RLU at 1.00% dilution.

10 The following PRO polypeptides tested positive in this assay: PRO535, PRO826, PRO819, PRO1126, PRO1160 and PRO1387.

EXAMPLE 155: Inhibitory Activity in Mixed Lymphocyte Reaction (MLR) Assay (Assay 67)

15 This example shows that one or more of the polypeptides of the invention are active as inhibitors of the proliferation of stimulated T-lymphocytes. Compounds which inhibit proliferation of lymphocytes are useful therapeutically where suppression of an immune response is beneficial.

The basic protocol for this assay is described in Current Protocols in Immunology, unit 3.12; edited by J E Coligan, A M Kruisbeek, D H Marglies, E M Shevach, W Strober, National Institutes of Health, Published by John Wiley & Sons, Inc.

20 More specifically, in one assay variant, peripheral blood mononuclear cells (PBMC) are isolated from mammalian individuals, for example a human volunteer, by leukopheresis (one donor will supply stimulator PBMCs, the other donor will supply responder PBMCs). If desired, the cells are frozen in fetal bovine serum and DMSO after isolation. Frozen cells may be thawed overnight in assay media (37°C, 5% CO₂) and then washed and resuspended to 3x10⁶ cells/ml of assay media (RPMI; 10% fetal bovine serum, 1% penicillin/streptomycin, 1% glutamine, 1% HEPES, 1% non-essential amino acids, 1% pyruvate). The stimulator PBMCs are prepared by irradiating the cells (about 3000 Rads).

The assay is prepared by plating in triplicate wells a mixture of:

100:1 of test sample diluted to 1% or to 0.1%,

50 :1 of irradiated stimulator cells, and

30 50 :1 of responder PBMC cells.

100 microliters of cell culture media or 100 microliter of CD4-IgG is used as the control. The wells are then incubated at 37°C, 5% CO₂ for 4 days. On day 5, each well is pulsed with tritiated thymidine (1.0 mCi/well; Amersham). After 6 hours the cells are washed 3 times and then the uptake of the label is evaluated.

35 In another variant of this assay, PBMCs are isolated from the spleens of Balb/c mice and C57B6 mice. The cells are teased from freshly harvested spleens in assay media (RPMI; 10% fetal bovine serum, 1% penicillin/streptomycin, 1% glutamine, 1% HEPES, 1% non-essential amino acids, 1% pyruvate) and the PBMCs are isolated by overlaying these cells over Lympholyte M (Organon Teknika), centrifuging at 2000

rpm for 20 minutes, collecting and washing the mononuclear cell layer in assay media and resuspending the cells to 1×10^7 cells/ml of assay media. The assay is then conducted as described above.

Any decreases below control is considered to be a positive result for an inhibitory compound, with decreases of less than or equal to 80% being preferred. However, any value less than control indicates an inhibitory effect for the test protein.

5 The following polypeptide tested positive in this assay: PRO1114, PRO836, PRO1159, PRO1312, PRO1192, PRO1195 and PRO1387.

EXAMPLE 156: Mouse Kidney Mesangial Cell Proliferation Assay (Assay 92)

10 This assay shows that certain polypeptides of the invention act to induce proliferation of mammalian kidney mesangial cells and, therefore, are useful for treating kidney disorders associated with decreased mesangial cell function such as Berger disease or other nephropathies associated with Schönlein-Henoch purpura,

celiac disease, dermatitis herpetiformis or Crohn disease. The assay is performed as follows. On day one, mouse kidney mesangial cells are plated on a 96 well plate in growth media (3:1 mixture of Dulbecco's modified Eagle's medium and Ham's F12 medium, 95% fetal bovine serum, 5% supplemented with 14 mM HEPES) and grown overnight. On day 2, PRO polypeptides are diluted at 2 concentrations (1% and 0.1%) in serum-free medium and added to the cells. Control samples are serum-free medium alone. On day 4, 20 μ l of the Cell Titer 96 Aqueous one solution reagent (Progenia) was added to each well and the colorimetric reaction was allowed to proceed for 2 hours. The absorbance (OD) is then measured at 490 nm. A positive in the assay is anything that gives an absorbance reading which is at least 15% above the control reading.

20 The following polypeptide tested positive in this assay: PRO819, PRO813 and PRO1066.

EXAMPLE 157: Pericyte c-Fos Induction (Assay 93)

25 This assay shows that certain polypeptides of the invention act to induce the expression of c-fos in pericyte cells and, therefore, are useful not only as diagnostic markers for particular types of pericyte-associated tumors but also for giving rise to antagonists which would be expected to be useful for the therapeutic treatment of pericyte-associated tumors. Specifically, on day 1, pericytes are received from VEC Technologies and all but 5 ml of media is removed from flask. On day 2, the pericytes are trypsinized, washed, spun and then plated onto 96 well plates. On day 7, the media is removed and the pericytes are treated with 100 μ l of PRO polypeptide test samples and controls (positive control = DME+5% serum +/- PDGF at 500 ng/ml; negative control = protein 32). Replicates are averaged and SD/CV are determined. Fold increase over Protein 32 (buffer control) value indicated by chemiluminescence units (RLU) luminometer reading verses frequency is plotted on a histogram. Two-fold above Protein 32 value is considered positive for the assay. ASY Matrix: Growth media = low glucose DMEM = 20% FBS + 1X pen strep + 1X fungizone. Assay Media = low glucose DMEM +5% FBS.

35 The following polypeptides tested positive in this assay: PRO943 and PRO819.

EXAMPLE 158: Detection of PRO Polypeptides That Affect Glucose or FFA Uptake by Primary Rat Adipocytes (Assay 94)

This assay is designed to determine whether PRO polypeptides show the ability to affect glucose or FFA uptake by adipocyte cells. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of disorders where either the stimulation or inhibition of glucose uptake by adipocytes would be beneficial including, for example, obesity, diabetes or hyper- or hypo-insulinemia.

In a 96 well format, PRO polypeptides to be assayed are added to primary rat adipocytes, and allowed to incubate overnight. Samples are taken at 4 and 16 hours and assayed for glycerol, glucose and FFA uptake. After the 16 hour incubation, insulin is added to the media and allowed to incubate for 4 hours. At this time, a sample is taken and glycerol, glucose and FFA uptake is measured. Media containing insulin without the PRO polypeptide is used as a positive reference control. As the PRO polypeptide being tested may either stimulate or inhibit glucose and FFA uptake, results are scored as positive in the assay if greater than 1.5 times or less than 0.5 times the insulin control.

The following PRO polypeptides tested positive as stimulators of glucose and/or FFA uptake in this assay: PRO1114, PRO1007, PRO1066, PRO848, PRO1182, PRO1198, PRO1192, PRO1271, PRO1375 and PRO1387.

The following PRO polypeptides tested positive as inhibitors of glucose and/or FFA uptake in this assay: PRO1184, PRO1360, PRO1309, PRO1154, PRO1181, PRO1186, PRO1160 and PRO1384.

EXAMPLE 159: Chondrocyte Re-differentiation Assay (Assay 110)

This assay shows that certain polypeptides of the invention act to induce redifferentiation of chondrocytes, therefore, are expected to be useful for the treatment of various bone and/or cartilage disorders such as, for example, sports injuries and arthritis. The assay is performed as follows. Porcine chondrocytes are isolated by overnight collagenase digestion of articular cartilage of metacarpophalangeal joints of 4-6 month old female pigs. The isolated cells are then seeded at 25,000 cells/cm² in Ham F-12 containing 10% FBS and 4 µg/ml gentamycin. The culture media is changed every third day and the cells are then seeded in 96 well plates at 5,000 cells/well in 100 µl of the same media without serum and 100 µl of the test PRO polypeptide, 5 nM staurosporin (positive control) or medium alone (negative control) is added to give a final volume of 200 µl/well. After 5 days of incubation at 37°C, a picture of each well is taken and the differentiation state of the chondrocytes is determined. A positive result in the assay occurs when the redifferentiation of the chondrocytes is determined to be more similar to the positive control than the negative control.

The following polypeptide tested positive in this assay: PRO1282, PRO1310, PRO619, PRO943, PRO820, PRO1080, PRO1016, PRO1007, PRO1056, PRO791, PRO1111, PRO1184, PRO1360, PRO1309, PRO1107, PRO1132, PRO1131, PRO848, PRO1181, PRO1186, PRO1159, PRO1312, PRO1192 and PRO1384.

EXAMPLE 160: Chondrocyte Proliferation Assay (Assay 111)

This assay is designed to determine whether PRO polypeptides of the present invention show the ability to induce the proliferation and/or redifferentiation of chondrocytes in culture. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of various bone and/or cartilage disorders such as, for example, sports injuries and arthritis.

Porcine chondrocytes are isolated by overnight collagenase digestion of articular cartilage of the metacarpophalangeal joint of 4-6 month old female pigs. The isolated cells are then seeded at 25,000 cells/cm² in Ham F-12 containing 10% FBS and 4 µg/ml gentamycin. The culture media is changed every third day and the cells are reseeded to 25,000 cells/cm² every five days. On day 12, the cells are seeded in 96 well plates at 5,000 cells/well in 100 µl of the same media without serum and 100 µl of either serum-free medium (negative control), staurosporin (final concentration of 5 nM; positive control) or the test PRO polypeptide are added to give a final volume of 200 µl/well. After 5 days at 37°C, 20 µl of Alamar blue is added to each well and the plates are incubated for an additional 3 hours at 37°C. The fluorescence is then measured in each well (Ex:530 nm; Em: 590 nm). The fluorescence of a plate containing 200 µl of the serum-free medium is measured to obtain the background. A positive result in the assay is obtained when the fluorescence of the PRO polypeptide treated sample is more like that of the positive control than the negative control.

The following PRO polypeptides tested positive in this assay: PRO1310, PRO844, PRO1312, PRO1192 and PRO1387.

EXAMPLE 161: Induction of Pancreatic β-Cell Precursor Proliferation (Assay 117)

This assay shows that certain polypeptides of the invention act to induce an increase in the number of pancreatic β-cell precursor cells and, therefore, are useful for treating various insulin deficient states in mammals, including diabetes mellitus. The assay is performed as follows. The assay uses a primary culture of mouse fetal pancreatic cells and the primary readout is an alteration in the expression of markers that represent either β-cell precursors or mature β-cells. Marker expression is measured by real time quantitative PCR (RTQ-PCR); wherein the marker being evaluated is a transcription factor called Pdx1.

The pancreata are dissected from E14 embryos (CD1 mice). The pancreata are then digested with collagenase/dispase in F12/DMEM at 37°C for 40 to 60 minutes (collagenase/dispase, 1.37 mg/ml, Boehringer Mannheim, #1097113). The digestion is then neutralized with an equal volume of 5% BSA and the cells are washed once with RPMI1640. At day 1, the cells are seeded into 12-well tissue culture plates (pre-coated with laminin, 20 µg/ml in PBS, Boehringer Mannheim, #124317). Cells from pancreata from 1-2 embryos are distributed per well. The culture medium for this primary culture is 14F/1640. At day 2, the media is removed and the attached cells washed with RPMI/1640. Two mls of minimal media are added in addition to the protein to be tested. At day 4, the media is removed and RNA prepared from the cells and marker expression analyzed by real time quantitative RT-PCR. A protein is considered to be active in the assay if it increases the expression of the relevant β-cell marker as compared to untreated controls.

14F/1640 is RPMI1640 (Gibco) plus the following:

group A 1:1000

group B 1:1000

recombinant human insulin 10 $\mu\text{g/ml}$

Aprotinin (50 $\mu\text{g/ml}$) 1:2000 (Boehringer manheim #981532)

Bovine pituitary extract (BPE) 60 $\mu\text{g/ml}$

Gentamycin 100 ng/ml

Group A : (in 10ml PBS)

Transferrin, 100mg (Sigma T2252)

Epidermal Growth Factor, 100 μg (BRL 100004)

Triiodothyronine, 10 μl of 5×10^{-6} M (Sigma T5516)

Ethanolamine, 100 μl of 10^{-1} M (Sigma E0135)

Phosphoethalamine, 100 μl of 10^{-1} M (Sigma P0503)

Selenium, 4 μl of 10^{-1} M (Aesar #12574)

Group C : (in 10ml 100% ethanol)

Hydrocortisone, 2 μl of 5×10^{-3} M (Sigma #H0135)

Progesterone, 100 μl of 1×10^{-3} M (Sigma #P6149)

Forskolin, 500 μl of 20mM (Calbiochem #344270)

Minimal media:

RPMI 1640 plus transferrin (10 $\mu\text{g/ml}$), insulin (1 $\mu\text{g/ml}$), gentamycin (100 ng/ml), aprotinin (50 $\mu\text{g/ml}$) and BPE (15 $\mu\text{g/ml}$).

Defined media:

RPMI 1640 plus transferrin (10 $\mu\text{g/ml}$), insulin (1 $\mu\text{g/ml}$), gentamycin (100 ng/ml) and aprotinin (50 $\mu\text{g/ml}$).

The following polypeptides tested positive in this assay: PRO1310, PRO1188, PRO1131 and PRO1387.

EXAMPLE 162: Detection of Polypeptides That Affect Glucose or FFA Uptake in Skeletal Muscle (Assay 106)

This assay is designed to determine whether PRO polypeptides show the ability to affect glucose or FFA uptake by skeletal muscle cells. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of disorders where either the stimulation or inhibition of glucose uptake by skeletal muscle would be beneficial including, for example, diabetes or hyper- or hypo-insulinemia.

In a 96 well format, PRO polypeptides to be assayed are added to primary rat differentiated skeletal muscle, and allowed to incubate overnight. Then fresh media with the PRO polypeptide and +/- insulin are added to the wells. The sample media is then monitored to determine glucose and FFA uptake by the skeletal muscle cells. The insulin will stimulate glucose and FFA uptake by the skeletal muscle, and insulin in media without the PRO polypeptide is used as a positive control, and a limit for scoring. As the PRO polypeptide being tested may either stimulate or inhibit glucose and FFA uptake, results are scored as positive in the assay

if greater than 1.5 times or less than 0.5 times the insulin control.

The following PRO polypeptides tested positive as either stimulators or inhibitors of glucose and/or FFA uptake in this assay: PRO358, PRO1016, PRO1007, PRO826, PRO1066, PRO1029 and PRO1309.

EXAMPLE 163: Fetal Hemoglobin Induction in an Erythroblastic Cell Line (Assay 107)

This assay is useful for screening PRO polypeptides for the ability to induce the switch from adult hemoglobin to fetal hemoglobin in an erythroblastic cell line. Molecules testing positive in this assay are expected to be useful for therapeutically treating various mammalian hemoglobin-associated disorders such as the various thalassemias. The assay is performed as follows. Erythroblastic cells are plated in standard growth medium at 1000 cells/well in a 96 well format. PRO polypeptides are added to the growth medium at a concentration of 0.2% or 2% and the cells are incubated for 5 days at 37°C. As a positive control, cells are treated with 100 μ M hemin and as a negative control, the cells are untreated. After 5 days, cell lysates are prepared and analyzed for the expression of gamma globin (a fetal marker). A positive in the assay is a gamma globin level at least 2-fold above the negative control.

The following polypeptides tested positive in this assay: PRO1114, PRO826, PRO1066, PRO844, PRO1192 and PRO1358.

EXAMPLE 164: Induction of Pancreatic β -Cell Precursor Differentiation (Assay 89)

This assay shows that certain polypeptides of the invention act to induce differentiation of pancreatic β -cell precursor cells into mature pancreatic β -cells and, therefore, are useful for treating various insulin deficient states in mammals, including diabetes mellitus. The assay is performed as follows. The assay uses a primary culture of mouse fetal pancreatic cells and the primary readout is an alteration in the expression of markers that represent either β -cell precursors or mature β -cells. Marker expression is measured by real time quantitative PCR (RTQ-PCR); wherein the marker being evaluated is insulin.

The pancreata are dissected from E14 embryos (CD1 mice). The pancreata are then digested with collagenase/dispase in F12/DMEM at 37°C for 40 to 60 minutes (collagenase/dispase, 1.37 mg/ml, Boehringer Mannheim, #1097113). The digestion is then neutralized with an equal volume of 5% BSA and the cells are washed once with RPMI1640. At day 1, the cells are seeded into 12-well tissue culture plates (pre-coated with laminin, 20 μ g/ml in PBS, Boehringer Mannheim, #124317). Cells from pancreata from 1-2 embryos are distributed per well. The culture medium for this primary culture is 14F/1640. At day 2, the media is removed and the attached cells washed with RPMI/1640. Two mls of minimal media are added in addition to the protein to be tested. At day 4, the media is removed and RNA prepared from the cells and marker expression analyzed by real time quantitative RT-PCR. A protein is considered to be active in the assay if it increases the expression of the relevant β -cell marker as compared to untreated controls.

14F/1640 is RPMI1640 (Gibco) plus the following:

group A 1:1000

group B 1:1000

recombinant human insulin 10 μ g/ml

Aprotinin (50 μ g/ml) 1:2000 (Boehringer manheim #981532)

Bovine pituitary extract (BPE) 60 μ g/ml

Gentamycin 100 ng/ml

Group A : (in 10ml PBS)

Transferrin, 100mg (Sigma T2252)

5 Epidermal Growth Factor, 100 μ g (BRL 100004)

Triiodothyronine, 10 μ l of 5x10⁻⁶ M (Sigma T5516)

Ethanolamine, 100 μ l of 10⁻¹ M (Sigma E0135)

Phosphoethalamine, 100 μ l of 10⁻¹ M (Sigma P0503)

Selenium, 4 μ l of 10⁻¹ M (Aesar #12574)

10 Group C : (in 10ml 100% ethanol)

Hydrocortisone, 2 μ l of 5X10⁻³ M (Sigma #H0135)

Progesterone, 100 μ l of 1X10⁻³ M (Sigma #P6149)

Forskolin, 500 μ l of 20mM (Calbiochem #344270)

Minimal media:

15 RPMI 1640 plus transferrin (10 μ g/ml), insulin (1 μ g/ml), gentamycin (100 ng/ml), aprotinin (50 μ g/ml) and BPE (15 μ g/ml).

Defined media:

RPMI 1640 plus transferrin (10 μ g/ml), insulin (1 μ g/ml), gentamycin (100 ng/ml) and aprotinin (50 μ g/ml).

20 The following polypeptides were positive in this assay: PRO1188, PRO1132, PRO1131 and PRO1181.

EXAMPLE 165: Skin Vascular Permeability Assay (Assay 64)

25 This assay shows that certain polypeptides of the invention stimulate an immune response and induce inflammation by inducing mononuclear cell, eosinophil and PMN infiltration at the site of injection of the animal. Compounds which stimulate an immune response are useful therapeutically where stimulation of an immune response is beneficial. This skin vascular permeability assay is conducted as follows. Hairless guinea pigs weighing 350 grams or more are anesthetized with ketamine (75-80 mg/Kg) and 5 mg/Kg xylazine intramuscularly (IM). A sample of purified polypeptide of the invention or a conditioned media test sample

30 is injected intradermally onto the backs of the test animals with 100 μ l per injection site. It is possible to have about 10-30, preferably about 16-24, injection sites per animal. One μ l of Evans blue dye (1% in physiologic buffered saline) is injected intracardially. Blemishes at the injection sites are then measured (mm diameter) at 1 hr and 6 hr post injection. Animals were sacrificed at 6 hrs after injection. Each skin injection site is biopsied and fixed in formalin. The skins are then prepared for histopathologic evaluation. Each site is

35 evaluated for inflammatory cell infiltration into the skin. Sites with visible inflammatory cell inflammation are scored as positive. Inflammatory cells may be neutrophilic, eosinophilic, monocytic or lymphocytic. At least a minimal perivascular infiltrate at the injection site is scored as positive, no infiltrate at the site of injection is

scored as negative.

The following polypeptide tested positive in this assay: PRO1007, PRO1358 and PRO1375.

EXAMPLE 166: Induction of Endothelial Cell Apoptosis (ELISA) (Assay 109)

The ability of PRO polypeptides to induce apoptosis in endothelial cells was tested in human venous umbilical vein endothelial cells (HUVEC, Cell Systems) using a 96-well format, in 0% serum media supplemented with 100 ng/ml VEGF, 0.1 % BSA, 1X penn/strep. A positive result in this assay indicates the usefulness of the polypeptide for therapeutically treating any of a variety of conditions associated with undesired endothelial cell growth including, for example, the inhibition of tumor growth. The 96-well plates used were manufactured by Falcon (No. 3072). Coating of 96 well plates were prepared by allowing gelatinization to occur for >30 minutes with 100 μ l of 0.2% gelatin in PBS solution. The gelatin mix was aspirated thoroughly before plating HUVEC cells at a final concentration of 2×10^4 cells/ml in 10% serum containing medium - 100 μ l volume per well. The cells were grown for 24 hours before adding test samples containing the PRO polypeptide of interest.

To all wells, 100 μ l of 0% serum media (Cell Systems) complemented with 100 ng/ml VEGF, 0.1 % BSA, 1X penn/strep was added. Test samples containing PRO polypeptides were added in triplicate at dilutions of 1%, 0.33% and 0.11%. Wells without cells were used as a blank and wells with cells only were used as a negative control. As a positive control, 1:3 serial dilutions of 50 μ l of a 3x stock of staurosporine were used. The cells were incubated for 24 to 35 hours prior to ELISA.

ELISA was used to determine levels of apoptosis preparing solutions according to the Boehringer Manual [Boehringer, Cell Death Detection ELISA plus, Cat No. 1 920 685]. Sample preparations: 96 well plates were spun down at 1 krpm for 10 minutes (200g); the supernatant was removed by fast inversion, placing the plate upside down on a paper towel to remove residual liquid. To each well, 200 μ l of 1X Lysis buffer was added and incubation allowed at room temperature for 30 minutes without shaking. The plates were spun down for 10 minutes at 1 krpm, and 20 μ l of the lysate (cytoplasmic fraction) was transferred into streptavidin coated MTP. 80 μ l of immunoreagent mix was added to the 20 μ l lysate in each well. The MTP was covered with adhesive foil and incubated at room temperature for 2 hours by placing it on an orbital shaker (200 rpm). After two hours, the supernatant was removed by suction and the wells rinsed three times with 250 μ l of 1X incubation buffer per well (removed by suction). Substrate solution was added (100 μ l) into each well and incubated on an orbital shaker at room temperature at 250 rpm until color development was sufficient for a photometric analysis (approx. after 10-20 minutes). A 96 well reader was used to read the plates at 405 nm, reference wavelength, 492 nm. The levels obtained for PIN 32 (control buffer) was set to 100%. Samples with levels > 130% were considered positive for induction of apoptosis.

The following PRO polypeptides tested positive in this assay: PRO844.

EXAMPLE 167: Guinea Pig Vascular Leak (Assay 32)

This assay is designed to determine whether PRO polypeptides of the present invention show the ability to induce vascular permeability. Polypeptides testing positive in this assay are expected to be useful

for the therapeutic treatment of conditions which would benefit from enhanced vascular permeability including, for example, conditions which may benefit from enhanced local immune system cell infiltration.

Hairless guinea pigs weighing 350 grams or more were anesthetized with Ketamine (75-80 mg/kg) and 5 mg/kg Xylazine intramuscularly. Test samples containing the PRO polypeptide or a physiological buffer without the test polypeptide are injected into skin on the back of the test animals with 100 μ l per injection site intradermally. There were approximately 16-24 injection sites per animal. One ml of Evans blue dye (1% in PBS) is then injected intracardially. Skin vascular permeability responses to the compounds (*i.e.*, blemishes at the injection sites of injection) are visually scored by measuring the diameter (in mm) of blue-colored leaks from the site of injection at 1, 6 and 24 hours post administration of the test materials. The mm diameter of blueness at the site of injection is observed and recorded as well as the severity of the vascular leakage. Blemishes of at least 5 mm in diameter are considered positive for the assay when testing purified proteins, being indicative of the ability to induce vascular leakage or permeability. A response greater than 7 mm diameter is considered positive for conditioned media samples. Human VEGF at 0.1 μ g/100 μ l is used as a positive control, inducing a response of 4-8 mm diameter.

The following PRO polypeptides tested positive in this assay: PRO1155.

EXAMPLE 168: Mouse Mesengial Cell Inhibition Assay (Assay 114)

This assay is designed to determine whether PRO polypeptides of the present invention show the ability to inhibit the proliferation of mouse mesengial cells in culture. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of such diseases or conditions where inhibition of mesengial cell proliferation would be beneficial such as, for example, cystic renal dysplasia, polycystic kidney disease, or other kidney disease associated with abnormal mesengial cell proliferation, renal tumors, and the like.

On day 1, mouse mesengial cells are plated on a 96 well plate in growth medium (a 3:1 mixture of Dulbecco's modified Eagle's medium and Ham's F12 medium, 95%; fetal bovine serum, 5%; supplemented with 14mM HEPES) and then are allowed to grow overnight. On day 2, the PRO polypeptide is diluted at 2 different concentrations (1%, 0.1%) in serum-free medium and is added to the cells. The negative control is growth medium without added PRO polypeptide. After the cells are allowed to incubate for 48 hours, 20 μ l of the Cell Titer 96 Aqueous one solution reagent (Promega) is added to each well and the colormetric reaction is allowed to proceed for 2 hours. The absorbance (OD) is then measured at 490 nm. A positive in the assay is an absorbance reading which is at least 10% above the negative control.

The following PRO polypeptides tested positive in this assay: PRO1192 and PRO1195.

Example 169: *In Vitro* Antitumor Assay (Assay 161)

The antiproliferative activity of various PRO polypeptides was determined in the investigational, disease-oriented *in vitro* anti-cancer drug discovery assay of the National Cancer Institute (NCI), using a sulforhodamine B (SRB) dye binding assay essentially as described by Skehan et al., J. Natl. Cancer Inst. 82:1107-1112 (1990). The 60 tumor cell lines employed in this study ("the NCI panel"), as well as conditions

for their maintenance and culture *in vitro* have been described by Monks et al., J. Natl. Cancer Inst. 83:757-766 (1991). The purpose of this screen is to initially evaluate the cytotoxic and/or cytostatic activity of the test compounds against different types of tumors (Monks et al., *supra*; Boyd, Cancer: Princ. Pract. Oncol. Update 3(10):1-12 [1989]).

Cells from approximately 60 human tumor cell lines were harvested with trypsin/EDTA (Gibco), washed once, resuspended in IMEM and their viability was determined. The cell suspensions were added by pipet (100 μ L volume) into separate 96-well microtiter plates. The cell density for the 6-day incubation was less than for the 2-day incubation to prevent overgrowth. Inoculates were allowed a preincubation period of 24 hours at 37°C for stabilization. Dilutions at twice the intended test concentration were added at time zero in 100 μ L aliquots to the microtiter plate wells (1:2 dilution). Test compounds were evaluated at five half-log dilutions (1000 to 100,000-fold). Incubations took place for two days and six days in a 5% CO₂ atmosphere and 100% humidity.

After incubation, the medium was removed and the cells were fixed in 0.1 ml of 10% trichloroacetic acid at 40°C. The plates were rinsed five times with deionized water, dried, stained for 30 minutes with 0.1 ml of 0.4% sulforhodamine B dye (Sigma) dissolved in 1% acetic acid, rinsed four times with 1% acetic acid to remove unbound dye, dried, and the stain was extracted for five minutes with 0.1 ml of 10 mM Tris base [tris(hydroxymethyl)aminomethane], pH 10.5. The absorbance (OD) of sulforhodamine B at 492 nm was measured using a computer-interfaced, 96-well microtiter plate reader.

A test sample is considered positive if it shows at least 50% growth inhibitory effect at one or more concentrations. The results are shown in the following table, where the abbreviations are as follows:

NSCL = non-small cell lung carcinoma

CNS = central nervous system

Table 7

<u>Test compound</u>	<u>Concentration</u>	<u>Days</u>	<u>Tumor Cell Line Type</u>	<u>C e l l L i n e Designation</u>
PRO1016	0.1 nM	2	Leukemia	K-568
PRO1016	0.1 nM	2	Leukemia	MOLT-4
PRO1016	0.1 nM	2	Leukemia	RPMI-8226
PRO1016	0.1 nM	2	NSCL	A549/ATCC
PRO1016	0.1 nM	2	NSCL	EKVX
PRO1016	0.1 nM	2	NSCL	NCI-H23
PRO1016	0.1 nM	2	NSCL	NCI-H522
PRO1016	0.1 nM	2	Colon	KM-12
PRO1016	0.1 nM	2	CNS	SF-295
PRO1016	0.1 nM	2	Melanoma	SK-MEL-5
PRO1016	0.1 nM	2	Melanoma	UACC-257
PRO1016	0.1 nM	2	Ovarian	OVCAR-3
PRO1016	0.1 nM	2	Ovarian	OVCAR-4
PRO1016	0.1 nM	2	Breast	NCI/SDR-RES
PRO1016	0.1 nM	2	Breast	T-47D
PRO1016	0.1 nM	6	Leukemia	CCRF-CEM
PRO1016	0.1 nM	6	Leukemia	K-562

Table 7 (cont')

	<u>Test compound</u>	<u>Concentration</u>	<u>Days</u>	<u>Tumor Cell Line Type</u>	<u>C e l l L i n e</u> <u>Designation</u>
5	PRO1016	0.1 nM	6	Leukemia	MOLT-4
	PRO1016	0.1 nM	6	Leukemia	RPMI-8226
	PRO1016	0.1 nM	6	NSCL	A549/ATCC
	PRO1016	0.1 nM	6	NSCL	EKVX
	PRO1016	0.1 nM	6	NSCL	HOP-62
10	PRO1016	0.1 nM	6	NSCL	NCI-H23
	PRO1016	0.1 nM	6	NSCL	NCI-H322M
	PRO1016	0.1 nM	6	NSCL	NCI-H460
	PRO1016	0.1 nM	6	NSCL	NCI-H522
	PRO1016	0.1 nM	6	Colon	COLO 205
15	PRO1016	0.1 nM	6	Colon	CHT-116
	PRO1016	0.1 nM	6	Colon	HCT-15
	PRO1016	0.1 nM	6	Colon	HT-29
	PRO1016	0.1 nM	6	Colon	SW-620
	PRO1016	0.1 nM	6	CNS	SF-295
20	PRO1016	0.1 nM	6	CNS	SF-539
	PRO1016	0.1 nM	6	CNS	SNB-19
	PRO1016	0.1 nM	6	CNS	U251
	PRO1016	0.1 nM	6	Melanoma	LOX IMVI
	PRO1016	0.1 nM	6	Melanoma	MALME-3M
25	PRO1016	0.1 nM	6	Melanoma	SK-MEL-28
	PRO1016	0.1 nM	6	Melanoma	SK-MEL-5
	PRO1016	0.1 nM	6	Melanoma	UACC-257
	PRO1016	0.1 nM	6	Melanoma	UACC-62
	PRO1016	0.1 nM	6	Ovarian	IGROV1
30	PRO1016	0.1 nM	6	Ovarian	OVCAR-3
	PRO1016	0.1 nM	6	Ovarian	OVCAR-4
	PRO1016	0.1 nM	6	Ovarian	OVCAR-8
	PRO1016	0.1 nM	6	Renal	ACHN
	PRO1016	0.1 nM	6	Renal	RXF 393
35	PRO1016	0.1 nM	6	Renal	SN12C
	PRO1016	0.1 nM	6	Renal	TK-10
	PRO1016	0.1 nM	6	Prostate	PC-3
	PRO1016	0.1 nM	6	Breast	MCF-7
	PRO1016	0.1 nM	6	Breast	NCI/ADR-RES
40	PRO1016	0.1 nM	6	Breast	MDA-MB-231
	PRO1016	0.1 nM	6	Breast	MDA-MB-435
	PRO1016	0.1 nM	6	Breast	MDA-N
	PRO1016	0.1 nM	6	Breast	BT-549
	PRO1016	0.1 nM	6	Breast	T-47D
45	PRO1186	95 nM	2	NSCL	NCI-H226
	PRO1186	95 nM	2	Colon	Colo205
	PRO1186	2.2 nM	6	Breast	MDA-N
	PRO1186	114 nM	2	NSCL	NCI-H322M
	PRO1186	114 nM	2	CNS	SF-268; SF-539
50	PRO1186	114 nM	2	Ovarian	IGFOV1
	PRO1186	114 nM	2	Renal	786-0; SN12C; TK-10
	PRO1186	114 nM	6	Leukemia	MOLT-4; RPMI-8226
	PRO1186	114 nM	6	Melanoma	LOX IMVI
	PRO1186	114 nM	6	Ovarian	OVCAR-4; SK-OV-3

Table 7 (cont')

	Test compound	Concentration	Days	Tumor Cell Line Type	C e l l L i n e Designation
	PRO1186	114 nM	6	Breast	MDA-MB-435; T-47D
5	PRO1186	8.1 nM	6	Leukemia	K-562
	PRO1186	8.1 nM	6	NSCL	HOP-62
	PRO1186	8.1 nM	6	Colon	Colo205; HCC-2998
	PRO1186	8.1 nM	6	Breast	T-47D
	PRO1186	15.4 nM	6	Leukemia	K-562
10	PRO1186	3.6 nM	2	Ovarian	OVCAR-3
	PRO1186	3.6 nM	6	NSCL	HOP-62

The results of these assays demonstrate that the positive testing PRO polypeptides are useful for inhibiting neoplastic growth in a number of different tumor cell types and may be used therapeutically therefor.. Antibodies against these PRO polypeptides are useful for affinity purification of these useful polypeptides. Nucleic acids encoding these PRO polypeptides are useful for the recombinant preparation of these polypeptides.

EXAMPLE 170: Gene Amplification in Tumors

This example shows that certain PRO polypeptide-encoding genes are amplified in the genome of certain human lung, colon and/or breast cancers and/or cell lines. Amplification is associated with overexpression of the gene product, indicating that the polypeptides are useful targets for therapeutic intervention in certain cancers such as colon, lung, breast and other cancers and diagnostic determination of the presence of those cancers. Therapeutic agents may take the form of antagonists of the PRO polypeptide, for example, murine-human chimeric, humanized or human antibodies against a PRO polypeptide.

The starting material for the screen was genomic DNA isolated from a variety cancers. The DNA is quantitated precisely, *e.g.*, fluorometrically. As a negative control, DNA was isolated from the cells of ten normal healthy individuals which was pooled and used as assay controls for the gene copy in healthy individuals (not shown). The 5' nuclease assay (for example, TaqMan™) and real-time quantitative PCR (for example, ABI Prizm 7700 Sequence Detection System™ (Perkin Elmer, Applied Biosystems Division, Foster City, CA)), were used to find genes potentially amplified in certain cancers. The results were used to determine whether the DNA encoding the PRO polypeptide is over-represented in any of the primary lung or colon cancers or cancer cell lines or breast cancer cell lines that were screened. The primary lung cancers were obtained from individuals with tumors of the type and stage as indicated in Table 8. An explanation of the abbreviations used for the designation of the primary tumors listed in Table 8 and the primary tumors and cell lines referred to throughout this example are given below.

The results of the TaqMan™ are reported in delta (Δ) Ct units. One unit corresponds to 1 PCR cycle or approximately a 2-fold amplification relative to normal, two units corresponds to 4-fold, 3 units to 8-fold amplification and so on. Quantitation was obtained using primers and a TaqMan™ fluorescent probe derived from the PRO polypeptide-encoding gene. Regions of the PRO polypeptide-encoding gene which are most likely to contain unique nucleic acid sequences and which are least likely to have spliced out introns are

preferred for the primer and probe derivation, *e.g.*, 3'-untranslated regions. The sequences for the primers and probes (forward, reverse and probe) used for the PRO polypeptide gene amplification analysis were as follows:

PRO290 (DNA35680-1212):

35680.tm.p:

5 5'-CCACCAATGGCAGCCCCACCT-3' (SEQ ID NO:428)

35680.tm.f:

5'-GACTGCCCTCCCTGCCA-3' (SEQ ID NO:429)

35680.tm.r:

10 5'-CAAAAAGCCTGGAAGTCTTCAAAG-3' (SEQ ID NO:430)

PRO341 (DNA26288-1239):

26288.tm.fl:

5'-CAGCTGGACTGCAGGTGCTA-3' (SEQ ID NO:431)

26288.tm.rl:

15 5'-CAGTGAGCACAGCAAGTGTCCCT-3' (SEQ ID NO:432)

26288.tm.pl:

5'-GGCCACCTCCTTGAGTCTTCAGTTCCT-3' (SEQ ID NO:433)

PRO535 (DNA49143-1429):

20 49143.tm.fl:

5'-CAACTACTGGCTAAAGCTGGTGAA-3' (SEQ ID NO:434)

49143.tm.rl:

5'-CCTTTCTGTATAGGTGATACCCAATGA-3' (SEQ ID NO:435)

49143.tm.pl:

25 5'-TGGCCATCCCTACCAGAGGCAAAA-3' (SEQ ID NO:436)

PRO619 (DNA49821-1562):

49821.tm.fl:

5'-CTGAAGACGACGCGGATTACTA-3' (SEQ ID NO:437)

30 49821.tm.rl:

5'-GGCAGAAATGGGAGGCAGA-3' (SEQ ID NO:438)

49821.tm.pl:

5'-TGCTCTGTTGGCTACGGCTTTAGTCCCTAG-3' (SEQ ID NO:439)

35 PRO809 (DNA57836-1338):

57836.tm.fl:

5'-AGCAGCAGCCATGTAGAATGAA-3' (SEQ ID NO:440)

57836.tm.r1:

5'-AATACGAACAGTGCACGCTGAT-3' (SEQ ID NO:441)

57836.tm.p1:

5'-TCCAGAGAGCCAAGCACGGCAGA-3' (SEQ ID NO:442)

5 PRO830 (DNA56866-1342):56866.tm.fl:

5'-TCTAGCCAGCTTGGCTCCAATA-3' (SEQ ID NO:443)

56866.tm.r1:

5'-CCTGGCTCTAGCACCAACTCATA-3' (SEQ ID NO:444)

10 56866.tm.p1:

5'-TCAGTGGCCCTAAGGAGATGGGCCT-3' (SEQ ID NO:445)

PRO848 (DNA59839-1461):59839.tm.fl:

15 5'-CAGGATACAGTGGGAATCTTGAGA-3' (SEQ ID NO:446)

59839.tm.r1:

5'-CCTGAAGGGCTTGGAGCTTAGT-3' (SEQ ID NO:447)

59839.tm.p1:

5'-TCTTTGGCCATTTCCCATGGCTCA-3' (SEQ ID NO:448)

20

PRO943 (DNA52192-1369):52192.tm.fl:

5'-CCCATGGCGAGGAGGAAT-3' (SEQ ID NO:449)

52192.tm.r1:

25 5'-TGC GTACGTGTGCCTTCAG-3' (SEQ ID NO:450)

52192.tm.p1:

5'-CAGCACCCCAGGCAGTCTGTGTGT-3' (SEQ ID NO:451)

PRO1005 (DNA57708-1411):30 57708.tm.fl:

5'-AACGTGCTACACGACCAGTGTACT-3' (SEQ ID NO:452)

57708.tm.r1:

5'-CACAGCATATTCAGATGACTAAATCCA-3' (SEQ ID NO:453)

57708.tm.p1:

35 5'-TTGTTTAGTTCTCCACCGTGTCTCCACAGAA-3' (SEQ ID NO:454)

PRO1009 (DNA57129-1413):57129.tm.fl:

5'-TGTCAGAATGCAACCTGGCTT-3' (SEQ ID NO:455)

57129.tm.rl:

5'-TGATGTGCCTGGCTCAGAAC-3' (SEQ ID NO:456)

5 57129.tm.pl:

5'-TGCACCTAGATGTCCCCAGCACCC-3' (SEQ ID NO:457)

PRO1097 (DNA59841-1460):59841.tm.fl:

10 5'-AAGATGCGCCAGGCTTCTTA-3' (SEQ ID NO:458)

59841.tm.rl:

5'-CTCCTGTACGGTCTGCTCACTTAT-3' (SEQ ID NO:459)

59841.tm.pl:

15 5'-TGGCTGTCAGTCCAGTGTGCATGG-3' (SEQ ID NO:460)

PRO1107 (DNA59606-1471):59606.tm.fl:

5'-GCATAGGGATAGATAAGATCCTGCTTTAT-3' (SEQ ID NO:461)

59606.tm.rl:

20 5'-CAAATTAAAGTACCCATCAGGAGAGAA-3' (SEQ ID NO:462)

59606.tm.pl:

5'-AAGTTGCTAAATATATACATTATCTGCGCCAAGTCCA-3' (SEQ ID NO:463)

PRO1111 (DNA58721-1475):25 58721.tm.fl:

5'-GTGCTGCCCACAATTCATGA-3' (SEQ ID NO:464)

58721.tm.rl:

5'-GTCCTTGGTATGGGTCTGAATTATAT-3' (SEQ ID NO:465)

58721.tm.pl:

30 5'-ACTCTCTGCACCCACAGTCACCACTATCTC-3' (SEQ ID NO:466)

PRO1153 (DNA59842-1502):59842.tm.fl:

5'-CTGAGGAACCAGCCATGTCTCT-3' (SEQ ID NO:467)

35 59842.tm.rl:

5'-GACCAGATGCAGGTACAGGATGA-3' (SEQ ID NO:468)

59842.tm.pl:

5'-CTGCCCCTTCAGTGATGCCAACCTT-3' (SEQ ID NO:469)

PRO1182 (DNA59848-1512):59848.tm.fl:

5 5'-GGGTGGAGGCTCACTGAGTAGA-3' (SEQ ID NO:470)

59848.tm.rl:

5'-CAATACAGGTAATGAAACTCTGCTTCTT-3' (SEQ ID NO:471)

59848.tm.pl:

10 5'-TCCTCTTAAGCATAGGCCATTTTCTCAGTTTAGACA-3' (SEQ ID NO:472)

PRO1184 (DNA59220-1514):59220.tm.fl:

5'-GGTGGTCTTGCTTGGTCTCAC-3' (SEQ ID NO:473)

59220.tm.rl:

15 5'-CCGTCGTTTCAGCAACATGAC-3' (SEQ ID NO:474)

59220.tm.pl:

5'-ACCGCCTACCGCTGTGCCCA-3' (SEQ ID NO:475)

PRO1187 (DNA62876-1517):

20 62876.tm.fl:

5'-CAGTAAAACCACAGGCTGGATTT-3' (SEQ ID NO:476)

62876.tm.rl:

5'-CCTGAGAGCAAGAAGGTTGAGAAT-3' (SEQ ID NO:477)

62876.tm.pl:

25 5'-TAGACAGGGACCATGGCCCGCA-3' (SEQ ID NO:478)

PRO1281 (DNA59820-1549):59820.tm.fl:

5'-TGGGCTGTAGAAGAGTTGTTG-3' (SEQ ID NO:479)

30 59820.tm.rl:

5'-TCCACACTTGGCCAGTTTAT-3' (SEQ ID NO:480)

59820.tm.pl:

5'-CCCAACTTCTCCCTTTTGGACCCT-3' (SEQ ID NO:481)

35 PRO1112 (DNA57702-1476):

57702.tm.fl

5'-GTCCCTTCACTGTTTAGAGCATGA-3' (SEQ ID NO:482)

57702.tm.p1

5'-ACTCTCCCCCTCAACAGCCTCCTGAG-3' (SEQ ID NO:483)

57702.tm.r1

5'-GTGGTCAGGGCAGATCCTTT-3' (SEQ ID NO:484)

5 PRO1185 (DNA62881-1515):62881.tm.fl:

5'-ACAGATCCAGGAGAGACTCCACA -3' (SEQ ID NO:485)

62881.tm.p1:

5'-AGCGGCGCTCCCAGCCTGAAT -3' (SEQ ID NO:486)

10 62881.tm.r1:

5'-CATGATTGGTCCTCAGTTCCATC -3' (SEQ ID NO:487)

PRO1245 (DNA64884-1527):64884.tm.fl:

15 5'-ATAGAGGGCTCCCAGAAGTG -3' (SEQ ID NO:488)

64884.tm.p1:

5'-CAGGGCCTTCAGGGCCTTCAC-3' (SEQ ID NO:489)

64884.tm.r1:

5'-GCTCAGCCAAACACTGTCA-3' (SEQ ID NO:490)

20 64884.tm.f2:

5'-GGGGCCCTGACAGTGTT -3' (SEQ ID NO:491)

64884.tm.p2:

5'-CTGAGCCGAGACTGGAGCATCTACAC-3' (SEQ ID NO:492)

64884.tm.r2:

25 5'-GTGGGCAGCGTCTTGTC-3' (SEQ ID NO:493)

The 5' nuclease assay reaction is a fluorescent PCR-based technique which makes use of the 5' exonuclease activity of Taq DNA polymerase enzyme to monitor amplification in real time. Two oligonucleotide primers (forward [.f] and reverse [.r]) are used to generate an amplicon typical of a PCR reaction. A third oligonucleotide, or probe (.p), is designed to detect nucleotide sequence located between the two PCR primers. The probe is non-extendible by Taq DNA polymerase enzyme, and is labeled with a reporter fluorescent dye and a quencher fluorescent dye. Any laser-induced emission from the reporter dye is quenched by the quenching dye when the two dyes are located close together as they are on the probe. During the amplification reaction, the Taq DNA polymerase enzyme cleaves the probe in a template-dependent manner. The resultant probe fragments disassociate in solution, and signal from the released reporter dye is free from the quenching effect of the second fluorophore. One molecule of reporter dye is liberated for each new molecule synthesized, and detection of the unquenched reporter dye provides the basis for quantitative interpretation of the data.

The 5' nuclease procedure is run on a real-time quantitative PCR device such as the ABI Prism 7700™ Sequence Detection. The system consists of a thermocycler, laser, charge-coupled device (CCD) camera and computer. The system amplifies samples in a 96-well format on a thermocycler. During amplification, laser-induced fluorescent signal is collected in real-time through fiber optics cables for all 96 wells, and detected at the CCD. The system includes software for running the instrument and for analyzing the data.

5' Nuclease assay data are initially expressed as Ct, or the threshold cycle. This is defined as the cycle at which the reporter signal accumulates above the background level of fluorescence. The ΔC_t values are used as quantitative measurement of the relative number of starting copies of a particular target sequence in a nucleic acid sample when comparing cancer DNA results to normal human DNA results.

Table 8 describes the stage, T stage and N stage of various primary tumors which were used to screen the PRO polypeptide compounds of the invention.

Table 8
Primary Lung and Colon Tumor Profiles

	<u>Primary Tumor Stage</u>	<u>Stage</u>	<u>Other Stage</u>	<u>Dukes Stage</u>	<u>T Stage</u>	<u>N Stage</u>
5	Human lung tumor AdenoCa (SRCC724) [LT1]	IIA			T1	N1
	Human lung tumor SqCCa (SRCC725) [LT1a]	IIB			T3	N0
	Human lung tumor AdenoCa (SRCC726) [LT2]	IB			T2	N0
	Human lung tumor AdenoCa (SRCC727) [LT3]	IIIA			T1	N2
	Human lung tumor AdenoCa (SRCC728) [LT4]	IB			T2	N0
	Human lung tumor SqCCa (SRCC729) [LT6]	IB			T2	N0
10	Human lung tumor Aden/SqCCa (SRCC730) [LT7]	IA			T1	N0
	Human lung tumor AdenoCa (SRCC731) [LT9]	IB			T2	N0
	Human lung tumor SqCCa (SRCC732) [LT10]	IIB			T2	N1
	Human lung tumor SqCCa (SRCC733) [LT11]	IIA			T1	N1
	Human lung tumor AdenoCa (SRCC734) [LT12]	IV			T2	N0
15	Human lung tumor AdenoSqCCa (SRCC735)[LT13]	IB			T2	N0
	Human lung tumor SqCCa (SRCC736) [LT15]	IB			T2	N0
	Human lung tumor SqCCa (SRCC737) [LT16]	IB			T2	N0
	Human lung tumor SqCCa (SRCC738) [LT17]	IIB			T2	N1
	Human lung tumor SqCCa (SRCC739) [LT18]	IB			T2	N0
20	Human lung tumor SqCCa (SRCC740) [LT19]	IB			T2	N0
	Human lung tumor LCCa (SRCC741) [LT21]	IIB			T3	N1
	Human lung AdenoCa (SRCC811) [LT22]	1A			T1	N0
	Human colon AdenoCa (SRCC742) [CT2]		M1	D	pT4	N0
	Human colon AdenoCa (SRCC743) [CT3]			B	pT3	N0
25	Human colon AdenoCa (SRCC744) [CT8]			B	T3	N0
	Human colon AdenoCa (SRCC745) [CT10]			A	pT2	N0
	Human colon AdenoCa (SRCC746) [CT12]		MO, R1	B	T3	N0
	Human colon AdenoCa (SRCC747) [CT14]		pMO, RO	B	pT3	pN0
	Human colon AdenoCa (SRCC748) [CT15]		M1, R2	D	T4	N2
30	Human colon AdenoCa (SRCC749) [CT16]		pMO	B	pT3	pN0
	Human colon AdenoCa (SRCC750) [CT17]			C1	pT3	pN1
	Human colon AdenoCa (SRCC751) [CT1]		MO, R1	B	pT3	N0
	Human colon AdenoCa (SRCC752) [CT4]			B	pT3	M0
	Human colon AdenoCa (SRCC753) [CT5]		G2	C1	pT3	pN0
35	Human colon AdenoCa (SRCC754) [CT6]		pMO, RO	B	pT3	pN0
	Human colon AdenoCa (SRCC755) [CT7]		G1	A	pT2	pN0
	Human colon AdenoCa (SRCC756) [CT9]		G3	D	pT4	pN2
	Human colon AdenoCa (SRCC757) [CT11]			B	T3	N0
40	Human colon AdenoCa (SRCC758) [CT18]		MO, RO	B	pT3	pN0

DNA Preparation:

DNA was prepared from cultured cell lines, primary tumors, normal human blood. The isolation was performed using purification kit, buffer set and protease and all from Quiagen, according to the manufacturer's instructions and the description below.

Cell culture lysis:

Cells were washed and trypsinized at a concentration of 7.5×10^8 per tip and pelleted by centrifuging at 1000 rpm for 5 minutes at 4°C, followed by washing again with 1/2 volume of PBS recentrifugation. The pellets were washed a third time, the suspended cells collected and washed 2x with PBS. The cells were then suspended into 10 ml PBS. Buffer C1 was equilibrated at 4°C. Qiagen protease #19155 was diluted into 6.25 ml cold ddH₂O to a final concentration of 20 mg/ml and equilibrated at 4°C. 10 ml of G2 Buffer was prepared

by diluting Qiagen RNase A stock (100 mg/ml) to a final concentration of 200 μ g/ml.

Buffer C1 (10 ml, 4°C) and ddH₂O (40 ml, 4°C) were then added to the 10 ml of cell suspension, mixed by inverting and incubated on ice for 10 minutes. The cell nuclei were pelleted by centrifuging in a Beckman swinging bucket rotor at 2500 rpm at 4°C for 15 minutes. The supernatant was discarded and the nuclei were suspended with a vortex into 2 ml Buffer C1 (at 4°C) and 6 ml ddH₂O, followed by a second 4°C centrifugation at 2500 rpm for 15 minutes. The nuclei were then resuspended into the residual buffer using 200 μ l per tip. G2 buffer (10 ml) was added to the suspended nuclei while gentle vortexing was applied. Upon completion of buffer addition, vigorous vortexing was applied for 30 seconds. Quiagen protease (200 μ l, prepared as indicated above) was added and incubated at 50°C for 60 minutes. The incubation and centrifugation was repeated until the lysates were clear (*e.g.*, incubating additional 30-60 minutes, pelleting at 3000 x g for 10 min., 4°C).

Solid human tumor sample preparation and lysis:

Tumor samples were weighed and placed into 50 ml conical tubes and held on ice. Processing was limited to no more than 250 mg tissue per preparation (1 tip/preparation). The protease solution was freshly prepared by diluting into 6.25 ml cold ddH₂O to a final concentration of 20 mg/ml and stored at 4°C. G2 buffer (20 ml) was prepared by diluting DNase A to a final concentration of 200 mg/ml (from 100 mg/ml stock). The tumor tissue was homogenated in 19 ml G2 buffer for 60 seconds using the large tip of the polytron in a laminar-flow TC hood in order to avoid inhalation of aerosols, and held at room temperature. Between samples, the polytron was cleaned by spinning at 2 x 30 seconds each in 2L ddH₂O, followed by G2 buffer (50 ml). If tissue was still present on the generator tip, the apparatus was disassembled and cleaned.

Quiagen protease (prepared as indicated above, 1.0 ml) was added, followed by vortexing and incubation at 50°C for 3 hours. The incubation and centrifugation was repeated until the lysates were clear (*e.g.*, incubating additional 30-60 minutes, pelleting at 3000 x g for 10 min., 4°C).

Human blood preparation and lysis:

Blood was drawn from healthy volunteers using standard infectious agent protocols and citrated into 10 ml samples per tip. Quiagen protease was freshly prepared by dilution into 6.25 ml cold ddH₂O to a final concentration of 20 mg/ml and stored at 4°C. G2 buffer was prepared by diluting RNase A to a final concentration of 200 μ g/ml from 100 mg/ml stock. The blood (10 ml) was placed into a 50 ml conical tube and 10 ml C1 buffer and 30 ml ddH₂O (both previously equilibrated to 4°C) were added, and the components mixed by inverting and held on ice for 10 minutes. The nuclei were pelleted with a Beckman swinging bucket rotor at 2500 rpm, 4°C for 15 minutes and the supernatant discarded. With a vortex, the nuclei were suspended into 2 ml C1 buffer (4°C) and 6 ml ddH₂O (4°C). Vortexing was repeated until the pellet was white. The nuclei were then suspended into the residual buffer using a 200 μ l tip. G2 buffer (10 ml) were added to the suspended nuclei while gently vortexing, followed by vigorous vortexing for 30 seconds. Quiagen protease was added (200 μ l) and incubated at 50°C for 60 minutes. The incubation and centrifugation was repeated until the lysates were clear (*e.g.*, incubating additional 30-60 minutes, pelleting at 3000 x g for 10 min., 4°C).

Purification of cleared lysates:(1) Isolation of genomic DNA:

Genomic DNA was equilibrated (1 sample per maxi tip preparation) with 10 ml QBT buffer. QF elution buffer was equilibrated at 50°C. The samples were vortexed for 30 seconds, then loaded onto equilibrated tips and drained by gravity. The tips were washed with 2 x 15 ml QC buffer. The DNA was eluted into 30 ml silanized, autoclaved 30 ml Corex tubes with 15 ml QF buffer (50°C). Isopropanol (10.5 ml) was added to each sample, the tubes covered with parafilm and mixed by repeated inversion until the DNA precipitated. Samples were pelleted by centrifugation in the SS-34 rotor at 15,000 rpm for 10 minutes at 4°C. The pellet location was marked, the supernatant discarded, and 10 ml 70% ethanol (4°C) was added. Samples were pelleted again by centrifugation on the SS-34 rotor at 10,000 rpm for 10 minutes at 4°C. The pellet location was marked and the supernatant discarded. The tubes were then placed on their side in a drying rack and dried 10 minutes at 37°C, taking care not to overdry the samples.

After drying, the pellets were dissolved into 1.0 ml TE (pH 8.5) and placed at 50°C for 1-2 hours. Samples were held overnight at 4°C as dissolution continued. The DNA solution was then transferred to 1.5 ml tubes with a 26 gauge needle on a tuberculin syringe. The transfer was repeated 5x in order to shear the DNA. Samples were then placed at 50°C for 1-2 hours.

(2) Quantitation of genomic DNA and preparation for gene amplification assay:

The DNA levels in each tube were quantified by standard A_{260} , A_{280} spectrophotometry on a 1:20 dilution (5 μ l DNA + 95 μ l ddH₂O) using the 0.1 ml quartz cuvetts in the Beckman DU640 spectrophotometer. A_{260}/A_{280} ratios were in the range of 1.8-1.9. Each DNA sample was then diluted further to approximately 200 ng/ml in TE (pH 8.5). If the original material was highly concentrated (about 700 ng/ μ l), the material was placed at 50°C for several hours until resuspended.

Fluorometric DNA quantitation was then performed on the diluted material (20-600 ng/ml) using the manufacturer's guidelines as modified below. This was accomplished by allowing a Hoeffer DyNA Quant 200 fluorometer to warm-up for about 15 minutes. The Hoechst dye working solution (#H33258, 10 μ l, prepared within 12 hours of use) was diluted into 100 ml 1 x TNE buffer. A 2 ml cuvette was filled with the fluorometer solution, placed into the machine, and the machine was zeroed. pGEM 3Zf(+) (2 μ l, lot #360851026) was added to 2 ml of fluorometer solution and calibrated at 200 units. An additional 2 μ l of pGEM 3Zf(+) DNA was then tested and the reading confirmed at 400 +/- 10 units. Each sample was then read at least in triplicate. When 3 samples were found to be within 10% of each other, their average was taken and this value was used as the quantification value.

The fluorometrically determined concentration was then used to dilute each sample to 10 ng/ μ l in ddH₂O. This was done simultaneously on all template samples for a single TaqMan plate assay, and with enough material to run 500-1000 assays. The samples were tested in triplicate with Taqman™ primers and probe both B-actin and GAPDH on a single plate with normal human DNA and no-template controls. The diluted samples were used provided that the CT value of normal human DNA subtracted from test DNA was +/- 1 Ct. The diluted, lot-qualified genomic DNA was stored in 1.0 ml aliquots at -80°C. Aliquots which were subsequently to be used in the gene amplification assay were stored at 4°C. Each 1 ml aliquot is enough

for 8-9 plates or 64 tests.

Gene amplification assay:

The PRO polypeptide compounds of the invention were screened in the following primary tumors and the resulting ΔC_t values greater than or equal to 1.0 are reported in Tables 9A-C below.

Table 9A

<u>ΔCt values in lung and colon primary tumors and cell line models</u>								
Primary	PRO290	PRO341	PRO535	PRO619	PRO1112	PRO809	PRO830	PRO848
<u>Tumor</u>								
LT-1a	—	—	—	—	—	—	1.13	—
LT3	—	—	—	1.04	—	—	—	—
				1.68				
LT7	—	—	—	1.21	—	—	—	—
				1.34				
LT9	—	—	—	1.19	—	—	—	—
				1.34				
LT10	—	—	—	1.41	1.135	—	—	—
				2.02				
LT11	1.63	—	1.40	1.69	1.525	1.40	1.25	1.04
				1.57				
LT12	—	—	—	1.81	1.195	1.61	1.35	1.22
LT13	1.47	—	1.37	2.13	1.635	1.03	—	—
				1.74				
LT15	1.67	—	—	2.08	1.775	—	—	—
				1.52				
LT16	—	1.12	—	—	—	—	—	—
LT17	1.22	1.33	1.42	1.83	1.455	1.10	1.17	—
				1.67				
LT18	—	—	—	1.32	1.255	—	—	—
				1.14				
LT19	2.07	—	—	2.33	—	—	1.31	—
				1.90				
LT21	—	1.15	—	1.15	—	1.05	—	1.07
				1.09				
CT2	1.56	—	—	1.22	2.265	—	—	—

Table 9A (cont')
 Δ Ct values in lung and colon primary tumors and cell line models

Primary Tumor	PRO290	PRO341	PRO535	PRO619	PRO1112	PRO809	PRO830	PRO848
CT3	---	---	1.28	1.49	---	---	---	---
CT8	---	---	---	---	1.065	---	---	---
CT10	---	---	1.34	---	1.575	---	---	---
CT12	---	---	---	---	1.315	---	---	---
CT14	---	---	1.29	---	1.895	---	---	---
CT15	---	---	1.10	1.00	1.465	---	---	---
CT16	---	---	1.35	1.02	1.255	---	---	---
CT17	---	---	1.26	1.23	---	---	---	---
CT1	---	---	---	1.12	1.245	---	---	---
CT4	---	---	1.03	1.25	1.535	---	---	---
CT5	---	---	---	1.34	1.975	---	---	---
CT6	---	---	1.00	1.06	1.575	---	---	---
CT11	1.16	---	1.25	1.80	2.285	---	---	---

Table 9B

ΔCt values in lung and colon primary tumors and cell line models

Primary Tumor	PRO943	PRO1005	PRO1009	PRO1185	PRO1245	PRO1097	PRO1107	PRO1111	PRO1153
LT-1	---	1.07	---	---	---	---	---	---	---
LT-1a	---	3.87	---	---	---	---	---	---	---
LT2	---	---	---	---	---	1.23	---	---	---
LT3	---	1.61	---	1.01	---	---	---	1.39	---
LT4	---	---	---	---	---	---	---	1.49	1.01
LT6	---	1.29	---	---	---	---	---	---	---
LT7	---	---	---	---	---	---	---	1.58	1.52
LT9	---	2.50	---	---	---	1.21	---	1.44	---
LT10	---	---	---	---	---	---	---	1.05	---
LT11	2.06	---	---	---	---	---	---	1.45	---
LT12	1.94	1.21	---	---	---	---	---	---	---
LT13	1.64	2.30	---	---	3.84	---	3.55	---	---
LT15	2.05	1.03	---	---	1.01	---	2.47	---	---
LT16	---	1.05	---	---	1.98	---	2.45	---	---
LT17	1.93	---	---	---	---	---	---	1.47	---
LT19	2.90	---	---	---	---	---	---	---	---
LT26	---	---	---	1.66	---	---	---	---	---
LT30	---	---	---	1.58	---	---	---	---	---
CT2	1.92	---	2.00	1.73	---	---	4.75	---	---
CT3	---	---	1.75	---	---	---	1.52	---	---
CT8	1.37	---	1.29	---	---	---	---	---	---
	1.12								

Table 9B (cont')
 Δ Ct values in lung and colon primary tumors and cell line models

Primary Tumor	PRO943	PRO1005	PRO1009	PRO1185	PRO1245	PRO1097	PRO1107	PRO1111	PRO1153
CT10	2.13	---	1.73	---	---	---	2.82	---	---
	1.67								
CT12	1.43	---	1.92	---	---	---	---	---	---
CT14	1.46	---	2.10	---	---	1.08	1.54	1.38	---
CT15	---	---	2.02	---	1.00	---	---	---	---
CT16	---	---	1.56	---	---	1.11	---	---	---
CT17	1.30	---	1.76	---	---	1.34	---	---	---
CT1	1.36	---	---	---	---	---	1.57	---	---
CT4	---	---	1.06	---	---	---	1.59	---	---
CT5	1.88	---	1.43	---	---	---	---	---	---
	2.51								
CT6	1.41	---	---	---	---	---	---	---	---
	1.75								
CT7	---	---	---	---	---	---	---	1.16	---
CT11	2.80	---	1.83	---	---	---	---	1.17	---
	2.61								
CT18	1.30	---	---	---	---	---	---	1.05	---
H522	---	---	---	---	1.10	---	---	---	---

Table 9C

 Δ Ct values in lung and colon primary tumors and cell line models

	Primary Tumor	PRO1182	PRO1184	PRO1187	PRO1281
5	LT-1	1.81	---	---	---
	LT-1a	---	1.14	---	---
			1.09		
10	LT4	1.43	1.37	---	---
			1.18		
	LT6	---	1.78	---	---
15			1.66		
			1.05		
	LT9	1.43	---	---	---
20	LT12	---	2.47	1.17	---
			2.61		
			1.80		
25	LT15	---	---	1.55	---
	LT16	---	1.01	1.33	---
	LT17	---	---	---	---
30	LT18	---	1.07	---	---
			1.13		
	LT19	---	1.19	---	---
			1.35		
			1.02		
	LT21	---	1.00	---	---
			1.20		
	CT2	---	---	---	1.15
	CT12	---	---	---	1.07

Because amplification of the various DNAs described above occurs in various cancerous tumors and tumor cell lines derived from various human tissues, these molecules likely play a significant role in tumor formation and/or growth. As a result, amplification and/or enhanced expression of these molecules can serve as a diagnostic for detecting the presence of tumor in an individual and antagonists (*e.g.*, antibodies) directed against the proteins encoded by the above described DNA molecules would be expected to have utility in cancer therapy.

EXAMPLE 171: Identification of Receptor/Ligand Interactions

In this assay, various PRO polypeptides are tested for ability to bind to a panel of potential receptor molecules for the purpose of identifying receptor/ligand interactions. The identification of a ligand for a known receptor, a receptor for a known ligand or a novel receptor/ligand pair is useful for a variety of indications including, for example, targeting bioactive molecules (linked to the ligand or receptor) to a cell known to express the receptor or ligand, use of the receptor or ligand as a reagent to detect the presence of the ligand or receptor in a composition suspected of containing the same, wherein the composition may comprise cells suspected of expressing the ligand or receptor, modulating the growth of or another biological or immunological activity of a cell known to express or respond to the receptor or ligand, modulating the immune response of cells or toward cells that express the receptor or ligand, allowing the preparation of agonists, antagonists and/or antibodies directed against the receptor or ligand which will modulate the growth of or a biological or immunological activity of a cell expressing the receptor or ligand, and various other indications which will be readily apparent to the ordinarily skilled artisan.

The assay is performed as follows. A PRO polypeptide of the present invention suspected of being a ligand for a receptor is expressed as a fusion protein containing the Fc domain of human IgG (an immunoadhesin). Receptor-ligand binding is detected by allowing interaction of the immunoadhesin polypeptide with cells (*e.g.* Cos cells) expressing candidate PRO polypeptide receptors and visualization of bound immunoadhesin with fluorescent reagents directed toward the Fc fusion domain and examination by microscope. Cells expressing candidate receptors are produced by transient transfection, in parallel, of defined subsets of a library of cDNA expression vectors encoding PRO polypeptides that may function as receptor molecules. Cells are then incubated for 1 hour in the presence of the PRO polypeptide immunoadhesin being tested for possible receptor binding. The cells are then washed and fixed with paraformaldehyde. The cells are then incubated with fluorescent conjugated antibody directed against the Fc portion of the PRO polypeptide immunoadhesin (*e.g.* FITC conjugated goat anti-human-Fc antibody). The cells are then washed again and examined by microscope. A positive interaction is judged by the presence of fluorescent labeling of cells transfected with cDNA encoding a particular PRO polypeptide receptor or pool of receptors and an absence of similar fluorescent labeling of similarly prepared cells that have been transfected with other cDNA or pools of cDNA. If a defined pool of cDNA expression vectors is judged to be positive for interaction with a PRO polypeptide immunoadhesin, the individual cDNA species that comprise the pool are tested individually (the pool is "broken down") to determine the specific cDNA that encodes a receptor able to interact with the PRO polypeptide immunoadhesin.

In another embodiment of this assay, an epitope-tagged potential ligand PRO polypeptide (e.g. histidine "His" tag) is allowed to interact with a panel of potential receptor PRO polypeptide molecules that have been expressed as fusions with the Fc domain of human IgG (immunoadhesins). Following a 1 hour co-incubation with the epitope tagged PRO polypeptide, the candidate receptors are each immunoprecipitated with protein A beads and the beads are washed. Potential ligand interaction is determined by western blot analysis of the immunoprecipitated complexes with antibody directed towards the epitope tag. An interaction is judged to occur if a band of the anticipated molecular weight of the epitope tagged protein is observed in the western blot analysis with a candidate receptor, but is not observed to occur with the other members of the panel of potential receptors.

Using these assays, the following receptor/ligand interactions have been herein identified:

- (1) PRO943 binds to FHF1, PRO183 (FHF2), PRO184 (FHF3) and PRO185 (FHF4) and vice versa.
- (2) PRO331 binds to PRO1133 and vice versa.
- (3) PRO363 binds to PRO1387 and vice versa.
- (4) PRO5723 binds to PRO1387 and vice versa.
- (5) PRO1114 binds to PRO3301 and PRO9940 and vice versa.
- (6) PRO9828 appears to be a novel fibroblast growth factor receptor (FGFR) ligand in that it binds to the known FGF receptors FGFR1, FGFR2IIIC, FGFR3IIIC and FGFR4. PRO9828 and agonists, therefore, will find use for activating the biological activities normally activated by FGF molecules including, for example, cell growth and proliferation. Antagonists of PRO9828 will find use in blocking the biological activities mediated through the FGF receptor.
- (7) PRO1181 binds to PRO7170, PRO361 and PRO846.

EXAMPLE 172: Tissue Expression Distribution

Oligonucleotide probes were constructed from the PRO polypeptide-encoding nucleotide sequences shown in the figures for use in quantitative PCR amplification reactions. The oligonucleotide probes were chosen so as to give an approximately 200-600 base pair amplified fragment from the 3' end of its associated template in a standard PCR reaction. The oligonucleotide probes were employed in standard quantitative PCR amplification reactions with cDNA libraries isolated from different human adult and/or fetal tissue sources and analyzed by agarose gel electrophoresis so as to obtain a quantitative determination of the level of expression of the PRO polypeptide-encoding nucleic acids in the various tissues tested. Knowledge of the expression pattern or the differential expression of the PRO polypeptide-encoding nucleic acids in various different human tissue types provides a diagnostic marker useful for tissue typing, with or without other tissue-specific markers, for determining the primary tissue source of a metastatic tumor, disease diagnosis, and the like. These assays provided the following results.

<u>DNA Molecule</u>	<u>Tissues w/ Significant Expression</u>	<u>Tissues w/o Significant Expression</u>
DNA16422-1209	substantia nigra, dendrocytes, uterus	hippocampus
DNA16435-1208	substantia nigra, dendrocytes, uterus	hippocampus
DNA26843-1389	dendrocytes, heart, uterus, colon tumor	hippocampus, substantia nigra, cartilage

	<u>DNA Molecule</u>	<u>Tissues w/ Significant Expression</u>	<u>Tissues w/o Significant Expression</u>
	DNA26844-1394	HUVEC, dendrocytes, cartilage	substantia nigra, hippocampus, uterus, prostate
5	DNA40621-1440	prostate, uterus, colon tumor	brain, heart, HUVEC, cartilage
	DNA44161-1434	colon tumor, dendrocytes	substantia nigra, hippocampus, prostate, uterus
	DNA44694-1500	dendrocytes, hippocampus, prostate	colon tumor, substantia nigra, heart
	DNA48320-1433	prostate, uterus	colon tumor, brain, heart, cartilage
	DNA49647-1398	brain, heart, prostate, uterus	cartilage
10	DNA53913-1490	hippocampus	substantia nigra, dendrocytes
	DNA53978-1443	dendrocytes, uterus, prostate	substantia nigra, colon tumor
	DNA53996-1442	spleen, prostate, uterus, hippocampus	substantia nigra, heart
	DNA56050-1455	prostate, uterus, cartilage, hippocampus	heart, colon tumor, dendrocytes
	DNA56110-1437	spleen, colon tumor, brain, prostate	heart
15	DNA56410-1414	uterus, dendrocytes	hippocampus, substantia nigra, heart
	DNA56436-1448	substantia nigra, prostate, hippocampus	dendrocytes, heart, HUVEC
	DNA56855-1447	prostate, uterus	brain, cartilage, heart, colon tumor
	DNA56860-1510	colon tumor	prostate, uterus, dendrocytes
	DNA56868-1478	colon tumor, prostate	uterus, brain, heart, cartilage
20	DNA56869-1545	prostate, uterus, cartilage	brain, colon tumor, spleen, heart
	DNA57699-1412	dendrocytes, hippocampus, prostate	substantia nigra, heart
	DNA57704-1452	brain, heart, spleen, uterus, prostate	colon tumor
	DNA57710-1451	dendrocytes, hippocampus, spleen, uterus	substantia nigra, heart
	DNA57711-1501	dendrocytes, hippocampus, heart, cartilage	substantia nigra
25	DNA57827-1493	colon tumor, hippocampus, prostate	substantia nigra, dendrocytes, uterus
	DNA58723-1588	substantia nigra, cartilage uterus	hippocampus, dendrocytes, HUVEC
	DNA58743-1609	brain, prostate, uterus	colon tumor, heart, spleen, cartilage
	DNA58846-1409	hippocampus, dendrocytes	substantia nigra, uterus, prostate, colon tumor
30	DNA58849-1494	prostate	brain, uterus, cartilage, heart, colon tumor
	DNA58850-1495	spleen, prostate, dendrocytes	hippocampus, substantia nigra, colon tumor
35	DNA59213-1487	spleen, cartilage, prostate, substantia nigra	heart, hippocampus, dendrocytes
	DNA59497-1496	dendrocytes, prostate, uterus, heart	cartilage, hippocampus, substantia nigra
	DNA59605-1418	dendrocytes, prostate, uterus	hippocampus, substantia nigra, colon tumor
	DNA59609-1470	dendrocytes	substantia nigra, hippocampus, heart, prostate, uterus, spleen
40	DNA59612-1466	prostate, dendrocytes	hippocampus, substantia nigra, uterus, colon tumor
	DNA59616-1465	dendrocytes, substantia nigra, colon tumor	hippocampus
	DNA59619-1464	dendrocytes, substantia nigra, colon tumor	hippocampus
45	DNA59625-1498	brain, colon tumor, prostate, uterus	THP-1 macrophages
	DNA59827-1426	substantia nigra, prostate, uterus	hippocampus, dendrocytes, heart
	DNA59828-1608	dendrocytes, substantia nigra, colon tumor	hippocampus
	DNA59853-1505	prostate	brain, uterus, spleen, heart, colon tumor
50	DNA59854-1459	cartilage	prostate, brain, heart, colon tumor
	DNA60283-1484	dendrocytes, spleen, prostate, uterus	hippocampus, substantia nigra, heart
	DNA60619-1482	dendrocytes, substantia nigra, colon tumor	hippocampus
	DNA60625-1507	cartilage	prostate, brain, heart, colon tumor
	DNA60629-1481	uterus, colon tumor, substantia nigra	hippocampus, dendrocytes, spleen, prostate
55	DNA61755-1554	dendrocytes, substantia nigra, colon tumor	hippocampus

<u>DNA Molecule</u>	<u>Tissues w/ Significant Expression</u>	<u>Tissues w/o Significant Expression</u>
DNA64852-1589	prostate, uterus	brain, heart, cartilage, colon tumor
DNA66308-1537	prostate, heart uterus	brain, colon tumor, cartilage
DNA68869-1610	spleen, prostate, heart, uterus, colon tumor, substantia nigra	hippocampus, dendrocytes, prostate

5

EXAMPLE 173: Isolation of cDNA Clones Encoding Human PRO846

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA39949. Based on the DNA39949 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO846.

10

Forward and reverse PCR primers were synthesized:

forward PCR primer 5'-CCCTGCAGTGCACCTACAGGGAAG-3' (SEQ ID NO:518)

reverse PCR primer 5'-CTGTCTTCCCCTGCTTGGCTGTGG-3' (SEQ ID NO:519)

15

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA39949 sequence which had the following nucleotide sequence

hybridization probe

5'-GGTGCAGGAAGGGTGGGATCCTCTTCTCTCGCTGCTCTGGCCACATC-3'

(SEQ ID NO:520)

20

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with one of the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO846 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

25

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO846 [herein designated as DNA44196-1353] (SEQ ID NO:516) and the derived protein sequence for PRO846.

30

The entire nucleotide sequence of UNQ422 (DNA44196-1353) is shown in Figure 329 (SEQ ID NO:516). Clone UNQ422 (DNA44196-1353) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 25-27 and ending at the stop codon at nucleotide positions 1021-1023 (Figure 329). The predicted polypeptide precursor is 332 amino acids long (Figure 330). The full-length PRO846 protein shown in Figure 330 has an estimated molecular weight of about 36,143 daltons and a pI of about 5.89. Important regions of the amino acid sequence of PRO846 include the signal peptide, the transmembrane domain, an N-glycosylation site, a sequence typical of fibrinogen beta and gamma chains C-terminal domain, and a sequence typical of Ig like V-type domain as shown in Figure 330. Clone UNQ422 (DNA44196-1353) has been deposited with ATCC and is assigned ATCC deposit no. 209847.

35

EXAMPLE 174: Isolation of cDNA Clones Encoding Human PRO363

A consensus sequence was obtained relative to a variety of EST sequences as described in Example 1 above, wherein the consensus sequence obtained is herein designated DNA42828. Based on the DNA42828 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO363.

A pair of PCR primers (forward and reverse) were synthesized:

forward PCR primer (42828.f1) 5'-CCAGTGCACAGCAGGCAACGAAGC-3' (SEQ ID NO:521)

reverse PCR primer (42828.r1) 5'-ACTAGGCTGTATGCCTGGGTGGGC-3' (SEQ ID NO:522)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA42828 sequence which had the following nucleotide sequence

hybridization probe (42828.p1)

5'-GTATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGC-3' (SEQ ID NO:523)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with the PCR primer pair identified above. A positive library was then used to isolate clones encoding the PRO363 gene using the probe oligonucleotide and one of the PCR primers. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue (LIB227).

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO363 [herein designated as UNQ318 (DNA45419-1252)] (SEQ ID NO:500) and the derived protein sequence for PRO363.

The entire nucleotide sequence of UNQ318 (DNA45419-1252) is shown in Figure 313 (SEQ ID NO:500). Clone UNQ318 (DNA45419-1252) contains a single open reading frame with an apparent translational initiation site at nucleotide positions 190-192 and ending at the stop codon at nucleotide positions 1309-1311 (Figure 313). The predicted polypeptide precursor is 373 amino acids long (Figure 314). The full-length PRO363 protein shown in Figure 314 has an estimated molecular weight of about 41,281 daltons and a pI of about 8.33. A transmembrane domain exists at amino acids 221 to 254 of the amino acid sequence shown in Figure 314 (SEQ ID NO:501). The PRO363 polypeptide also possesses at least two myelin P0 protein domains from about amino acids 15 to 56 and from about amino acids 87 to 116. Clone UNQ318 (DNA45419-1252) has been deposited with ATCC on February 5, 1998 and is assigned ATCC deposit no. 209616.

Analysis of the amino acid sequence of the full-length PRO363 polypeptide suggests that it possesses significant sequence similarity to the cell surface protein HCAR, thereby indicating that PRO363 may be a novel HCAR homolog. More specifically, an analysis of the Dayhoff database (version 35.45 SwissProt 35) evidenced significant homology between the PRO363 amino acid sequence and the following Dayhoff sequences, HS46KDA_1, HSU90716_1, MMCARH_1, MMCARHOM_1, MMU90715_1, A33_HUMAN, P_W14146, P_W14158, A42632 and B42632.

EXAMPLE 175: Isolation of cDNA Clones Encoding a Human PRO9828

A consensus DNA sequence was assembled relative to other nucleic sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA139814. Based on the DNA139814 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO9828. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., Current Protocols in Molecular Biology, *supra*, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

PCR primers (forward and reverse) were synthesized:

5'-AATCTCAGCACCAGCCACTCAGAGCA-3' (SEQ ID NO:524)

5'-GTAAAGAGGGTGCCCTTCCAGCGA-3' (SEQ ID NO:525)

5'-TATCCCAATGCCTCCCCACTGCTC-3' (SEQ ID NO:526)

5'-GATGAAGTTGGCGAAGGGGCGGCA-3' (SEQ ID NO:527)

RNA for construction of the cDNA libraries was isolated from human fetal liver tissue. The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for a full-length PRO9828 polypeptide (designated herein as DNA142238-2768 [Figure 323, SEQ ID NO:510]) and the derived protein sequence for that PRO9828 polypeptide.

The full length clone identified above contained a single open reading frame with an apparent translational initiation site at nucleotide positions 232-234 and a stop signal at nucleotide positions 985-987 (Figure 323, SEQ ID NO:510). The predicted polypeptide precursor is 251 amino acids long, has a calculated molecular weight of approximately 27,954 daltons and an estimated pI of approximately 9.22. Analysis of the full-length PRO9828 sequence shown in Figure 324 (SEQ ID NO:511) evidences the presence of a variety of important polypeptide domains as shown in Figure 324, wherein the locations given for those important polypeptide domains are approximate as described above. Chromosome mapping evidences that the PRO9828-encoding nucleic acid maps to chromosome 12p13 in humans. Clone DNA142238-2768 has been deposited with ATCC on October 5, 1999 and is assigned ATCC deposit no. 819-PTA.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using the ALIGN-2 sequence alignment analysis of the full-length sequence shown in Figure 324 (SEQ ID NO:511), evidenced sequence

identity between the PRO9828 amino acid sequence and the following Dayhoff sequences: P_Y08581, AB018122_1, FGF3_HUMAN, P_R70824, S54407, P_R80780, P_Y23761, P_W92312, OMFGF6_1 and P_R80871.

EXAMPLE 176: Isolation of cDNA Clones Encoding a Human PRO7170

5 DNA108722-2743 was identified by applying a proprietary signal sequence finding algorithm developed by Genentech, Inc. (South San Francisco, CA) upon ESTs as well as clustered and assembled EST fragments from public (e.g., Genbank) and/or private (LIFESEQ®, Incyte Pharmaceuticals, Inc., Palo Alto, CA) databases. The signal sequence algorithm computes a secretion signal score based on the character of the DNA nucleotides surrounding the first and optionally the second methionine codon(s) (ATG) at the 5'-end of the sequence or sequence fragment under consideration. The nucleotides following the first ATG must code for at least 35 unambiguous amino acids without any stop codons. If the first ATG has the required amino acids, the second is not examined. If neither meets the requirement, the candidate sequence is not scored. In order to determine whether the EST sequence contains an authentic signal sequence, the DNA and corresponding amino acid sequences surrounding the ATG codon are scored using a set of seven sensors (evaluation parameters) known to be associated with secretion signals.

15 Use of the above described signal sequence algorithm allowed identification of an EST cluster sequence from the LIFESEQ® database, Incyte Pharmaceuticals, Palo Alto, designated herein as CLU57836. This EST cluster sequence was then compared to a variety of expressed sequence tag (EST) databases which included public EST databases (e.g., Genbank) and a proprietary EST DNA database (LIFESEQ®, Incyte Pharmaceuticals, Palo Alto, CA) to identify existing homologies. The homology search was performed using the computer program BLAST or BLAST2 (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Those comparisons resulting in a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into a consensus DNA sequence with the program "phrap" (Phil Green, University of Washington, Seattle, Washington). The consensus sequence obtained therefrom is herein designated DNA58756.

25 In light of an observed sequence homology between the DNA58756 sequence and an EST sequence encompassed within clone no. 2251462 from the LIFESEQ® database, Incyte Pharmaceuticals, Palo Alto, CA, clone no. 2251462 was purchased and the cDNA insert was obtained and sequenced. It was found herein that that cDNA insert encoded a full-length protein. The sequence of this cDNA insert is shown in Figure 325 and is herein designated as DNA108722-2743.

30 Clone DNA108722-2743 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 60-62 and ending at the stop codon at nucleotide positions 1506-1508 (Figure 325). The predicted polypeptide precursor is 482 amino acids long (Figure 326). The full-length PRO7170 protein shown in Figure 326 has an estimated molecular weight of about 49,060 daltons and a pI of about 4.74. Analysis of the full-length PRO7170 sequence shown in Figure 326 (SEQ ID NO:513) evidences the presence of a variety of important polypeptide domains as shown in Figure 326, wherein the locations given for those important polypeptide domains are approximate as described above. Clone DNA108722-2743 has been

deposited with ATCC on August 17, 1999 and is assigned ATCC Deposit No. 552-PTA.

An analysis of the Dayhoff database (version 35.45 SwissProt 35), using the ALIGN-2 sequence alignment analysis of the full-length sequence shown in Figure 326 (SEQ ID NO:513), evidenced sequence identity between the PRO7170 amino acid sequence and the following Dayhoff sequences: P_Y12291, I47141, D88733_1, DMC56G7_1, P_Y11606, HWP1_CANAL, HSMUC5BEX_1, HSU78550_1, HSU70136_1, and SGS3_DROME.

EXAMPLE 177: Isolation of cDNA Clones Encoding Human PRO361

A consensus DNA sequence was assembled relative to other EST sequences using phrap as described in Example 1 above. This consensus sequence is herein designated DNA40654. Based on the DNA40654 consensus sequence, oligonucleotides were synthesized: 1) to identify by PCR a cDNA library that contained the sequence of interest, and 2) for use as probes to isolate a clone of the full-length coding sequence for PRO361.

Forward and reverse PCR primers were synthesized as follows:

<u>forward PCR primer</u>	5'-AGGGAGGATTATCCTTGACCTTTGAAGACC-3'	(SEQ ID NO:528)
<u>forward PCR primer</u>	5'-GAAGCAAGTGCCCAGCTC-3'	(SEQ ID NO:529)
<u>forward PCR primer</u>	5'-CGGGTCCCTGCTCTTTGG-3'	(SEQ ID NO:530)
<u>reverse PCR primer</u>	5'-CACCGTAGCTGGGAGCGCACTCAC-3'	(SEQ ID NO:531)
<u>reverse PCR primer</u>	5'-AGTGTAAGTCAAGCTCCC-3'	(SEQ ID NO:532)

Additionally, a synthetic oligonucleotide hybridization probe was constructed from the consensus DNA40654 sequence which had the following nucleotide sequence
hybridization probe

5'- GCTTCCTGACACTAAGGCTGTCTGCTAGTCAGAATTGCCTCAAAAAGAG-3' (SEQ ID NO:533)

In order to screen several libraries for a source of a full-length clone, DNA from the libraries was screened by PCR amplification with one of the PCR primer pairs identified above. A positive library was then used to isolate clones encoding the PRO361 gene using the probe oligonucleotide. RNA for construction of the cDNA libraries was isolated from human fetal kidney tissue.

DNA sequencing of the clones isolated as described above gave the full-length DNA sequence for PRO361 [herein designated as DNA45410-1250] (SEQ ID NO:514) and the derived protein sequence for PRO361.

The entire nucleotide sequence of DNA45410-1250 is shown in Figure 327 (SEQ ID NO:514). Clone DNA45410-1250 contains a single open reading frame with an apparent translational initiation site at nucleotide positions 226-228 and ending at the stop codon at nucleotide positions 1519-1521 (Figure 327). The predicted polypeptide precursor is 431 amino acids long (Figure 328). The full-length PRO361 protein shown in Figure 328 has an estimated molecular weight of about 46,810 daltons and a pI of about 6.45. In addition, regions of interest including the transmembrane domain (amino acids 380-409) and sequences typical of the arginase family of proteins (amino acids 3-14 and 39-57) are designated in Figure 328. Clone DNA45410-1250 has been deposited with ATCC and is assigned ATCC deposit no. ATCC 209621.

Analysis of the amino acid sequence of the full-length PRO361 polypeptide suggests that portions of it possess significant homology to the mucin and/or chitinase proteins, thereby indicating that PRO361 may be a novel mucin and/or chitinase protein.

EXAMPLE 178: Isolation of cDNA Clones Encoding a Human PRO183, PRO184, PRO185, PRO5723, PRO3301 or PRO9940

DNA molecules encoding the PRO183, PRO184, PRO185, PRO5723, PRO3301 or PRO9940 polypeptides shown in the accompanying figures were obtained through GenBank.

Deposit of Material

The following materials have been deposited with the American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209, USA (ATCC):

Table 10

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
15	DNA45410-1250	209621	February 5, 1998
	DNA108722-2743	552-PTA	August 17, 1999
	DNA142238-2768	819-PTA	October 5, 1999
	DNA40981-1234	209439	November 7, 1997
	DNA45419-1252	209616	February 5, 1998
20	DNA44196-1353	209847	May 6, 1998
	DNA16422-1209	209929	June 2, 1998
	DNA16435-1208	209930	June 2, 1998
	DNA21624-1391	209917	June 2, 1998
	DNA23334-1392	209918	June 2, 1998
25	DNA26288-1239	209792	April 21, 1998
	DNA26843-1389	203099	August 4, 1998
	DNA26844-1394	209926	June 2, 1998
	DNA30862-1396	209920	June 2, 1998
	DNA35680-1212	209790	April 21, 1998
30	DNA40621-1440	209922	June 2, 1998
	DNA44161-1434	209907	May 27, 1998
	DNA44694-1500	203114	August 11, 1998
	DNA45495-1550	203156	August 25, 1998
	DNA47361-1154	209431	November 7, 1997
35	DNA47394-1572	203109	August 11, 1998
	DNA48320-1433	209904	May 27, 1998
	DNA48334-1435	209924	June 2, 1998
	DNA48606-1479	203040	July 1, 1998
	DNA49141-1431	203003	June 23, 1998
40	DNA49142-1430	203002	June 23, 1998
	DNA49143-1429	203013	June 23, 1998
	DNA49647-1398	209919	June 2, 1998
	DNA49819-1439	209931	June 2, 1998
	DNA49820-1427	209932	June 2, 1998
45	DNA49821-1562	209981	June 16, 1998
	DNA52192-1369	203042	July 1, 1998
	DNA52598-1518	203107	August 11, 1998
	DNA53913-1490	203162	August 25, 1998

Table 10 (cont')

	DNA53978-1443	209983	June 16, 1998
	DNA53996-1442	209921	June 2, 1998
	DNA56041-1416	203012	June 23, 1998
	DNA56047-1456	209948	June 9, 1998
5	DNA56050-1455	203011	June 23, 1998
	DNA56110-1437	203113	August 11, 1998
	DNA56113-1378	203049	July 1, 1998
	DNA56410-1414	209923	June 2, 1998
	DNA56436-1448	209902	May 27, 1998
10	DNA56855-1447	203004	June 23, 1998
	DNA56859-1445	203019	June 23, 1998
	DNA56860-1510	209952	June 9, 1998
	DNA56865-1491	203022	June 23, 1998
	DNA56866-1342	203023	June 23, 1998
15	DNA56868-1209	203024	June 23, 1998
	DNA56869-1545	203161	August 25, 1998
	DNA56870-1492	209925	June 2, 1998
	DNA57033-1403	209905	May 27, 1998
	DNA57037-1444	209903	May 27, 1998
20	DNA57129-1413	209977	June 16, 1998
	DNA57690-1374	209950	June 9, 1998
	DNA57693-1424	203008	June 23, 1998
	DNA57694-1341	203017	June 23, 1998
	DNA57695-1340	203006	June 23, 1998
25	DNA57699-1412	203020	June 23, 1998
	DNA57702-1476	209951	June 9, 1998
	DNA57704-1452	209953	June 9, 1998
	DNA57708-1411	203021	June 23, 1998
	DNA57710-1451	203048	July 1, 1998
30	DNA57711-1501	203047	July 1, 1998
	DNA57827-1493	203045	July 1, 1998
	DNA57834-1339	209954	June 9, 1998
	DNA57836-1338	203025	June 23, 1998
	DNA57838-1337	203014	June 23, 1998
35	DNA57844-1410	203010	June 23, 1998
	DNA58721-1475	203110	August 11, 1998
	DNA58723-1588	203133	August 18, 1998
	DNA58737-1473	203136	August 18, 1998
	DNA58743-1609	203154	August 25, 1998
40	DNA58846-1409	209957	June 9, 1998
	DNA58848-1472	209955	June 9, 1998
	DNA58849-1494	209958	June 9, 1998
	DNA58850-1495	209956	June 9, 1998
	DNA58853-1423	203016	June 23, 1998
45	DNA58855-1422	203018	June 23, 1998
	DNA59205-1421	203009	June 23, 1998
	DNA59211-1450	209960	June 9, 1998
	DNA59213-1487	209959	June 9, 1998
	DNA59214-1449	203046	July 1, 1998
50	DNA59215-1425	209961	June 9, 1998
	DNA59220-1514	209962	June 9, 1998
	DNA59488-1603	203157	August 25, 1998
	DNA59493-1420	203050	July 1, 1998
	DNA59497-1496	209941	June 4, 1998
55	DNA59588-1571	203106	August 11, 1998

Table 10 (cont')

	DNA59603-1419	209944	June 9, 1998
	DNA59605-1418	203005	June 23, 1998
	DNA59606-1471	209945	June 9, 1998
	DNA59607-1497	209957	June 9, 1998
5	DNA59609-1470	209963	June 9, 1998
	DNA59610-1559	209990	June 16, 1998
	DNA59612-1466	209947	June 9, 1998
	DNA59613-1417	203007	June 23, 1998
	DNA59616-1465	209991	June 16, 1998
10	DNA59619-1464	203041	July 1, 1998
	DNA59620-1463	209989	June 16, 1998
	DNA59625-1498	209992	June 17, 1998
	DNA59767-1489	203108	August 11, 1998
	DNA59776-1600	203128	August 18, 1998
15	DNA59777-1480	203111	August 11, 1998
	DNA59820-1549	203129	August 18, 1998
	DNA59827-1426	203089	August 4, 1998
	DNA59828-1608	203158	August 25, 1998
	DNA59838-1462	209976	June 16, 1998
20	DNA59839-1461	209988	June 16, 1998
	DNA59841-1460	203044	July 1, 1998
	DNA59842-1502	209982	June 16, 1998
	DNA59846-1503	209978	June 16, 1998
	DNA59847-1511	203098	August 4, 1998
25	DNA59848-1512	203088	August 4, 1998
	DNA59849-1504	209986	June 16, 1998
	DNA59853-1505	209985	June 16, 1998
	DNA59854-1459	209974	June 16, 1998
	DNA60283-1484	203043	July 1, 1998
30	DNA60615-1483	209980	June 16, 1998
	DNA60619-1482	209993	June 16, 1998
	DNA60621-1516	203091	August 4, 1998
	DNA60622-1525	203090	August 4, 1998
	DNA60625-1507	209975	June 16, 1998
35	DNA60627-1508	203092	August 4, 1998
	DNA60629-1481	209979	June 16, 1998
	DNA61755-1554	203112	August 11, 1998
	DNA61873-1574	203132	August 18, 1998
	DNA62814-1521	203093	August 4, 1998
40	DNA62872-1509	203100	August 4, 1998
	DNA62876-1517	203095	August 4, 1998
	DNA62881-1515	203096	August 4, 1998
	DNA64852-1589	203127	August 18, 1998
	DNA64884-1527	203155	August 25, 1998
45	DNA64890-1612	203131	August 18, 1998
	DNA65412-1523	203094	August 4, 1998
	DNA66308-1537	203159	August 25, 1998
	DNA66309-1538	203235	September 15, 1998
	DNA67004-1614	203115	August 11, 1998
50	DNA68869-1610	203164	August 25, 1998
	DNA68872-1620	203160	August 25, 1998
	DNA71159-1617	203135	August 18, 1998

These deposit were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposits will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC §122 and the Commissioner's rules pursuant thereto (including 37 CFR §1.14 with particular reference to 886 OG 638).

The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. Isolated nucleic acid having at least 80% sequence identity to a nucleotide sequence that encodes a polypeptide comprising an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:6), Figure 6 (SEQ ID NO:8), Figure 9 (SEQ ID NO:14), Figure 12 (SEQ ID NO:20), Figure 15 (SEQ ID NO:23), Figure 18 (SEQ ID NO:28), Figure 20 (SEQ ID NO:30), Figure 23 (SEQ ID NO:33), Figure 25 (SEQ ID NO:36), Figure 27 (SEQ ID NO:41), Figure 30 (SEQ ID NO:47), Figure 32 (SEQ ID NO:52), Figure 34 (SEQ ID NO:57), Figure 36 (SEQ ID NO:62), Figure 38 (SEQ ID NO:67), Figure 41 (SEQ ID NO:73), Figure 47 (SEQ ID NO:84), Figure 49 (SEQ ID NO:95), Figure 51 (SEQ ID NO:97), Figure 53 (SEQ ID NO:99), Figure 57 (SEQ ID NO:103), Figure 64 (SEQ ID NO:113), Figure 66 (SEQ ID NO:115), Figure 68 (SEQ ID NO:117), Figure 70 (SEQ ID NO:119), Figure 72 (SEQ ID NO:124), Figure 74 (SEQ ID NO:129), Figure 76 (SEQ ID NO:135), Figure 79 (SEQ ID NO:138), Figure 83 (SEQ ID NO:146), Figure 85 (SEQ ID NO:148), Figure 88 (SEQ ID NO:151), Figure 90 (SEQ ID NO:153), Figure 93 (SEQ ID NO:156), Figure 95 (SEQ ID NO:158), Figure 97 (SEQ ID NO:160), Figure 99 (SEQ ID NO:165), Figure 101 (SEQ ID NO:167), Figure 103 (SEQ ID NO:169), Figure 105 (SEQ ID NO:171), Figure 109 (SEQ ID NO:175), Figure 111 (SEQ ID NO:177), Figure 113 (SEQ ID NO:179), Figure 115 (SEQ ID NO:181), Figure 117 (SEQ ID NO:183), Figure 120 (SEQ ID NO:189), Figure 122 (SEQ ID NO:194), Figure 125 (SEQ ID NO:197), Figure 127 (SEQ ID NO:199), Figure 129 (SEQ ID NO:201), Figure 131 (SEQ ID NO:203), Figure 133 (SEQ ID NO:205), Figure 135 (SEQ ID NO:207), Figure 137 (SEQ ID NO:209), Figure 139 (SEQ ID NO:211), Figure 141 (SEQ ID NO:213), Figure 144 (SEQ ID NO:216), Figure 147 (SEQ ID NO:219), Figure 149 (SEQ ID NO:221), Figure 151 (SEQ ID NO:223), Figure 153 (SEQ ID NO:225), Figure 155 (SEQ ID NO:227), Figure 157 (SEQ ID NO:229), Figure 159 (SEQ ID NO:231), Figure 161 (SEQ ID NO:236), Figure 163 (SEQ ID NO:241), Figure 165 (SEQ ID NO:246), Figure 167 (SEQ ID NO:248), Figure 169 (SEQ ID NO:250), Figure 171 (SEQ ID NO:253), Figure 174 (SEQ ID NO:256), Figure 176 (SEQ ID NO:258), Figure 178 (SEQ ID NO:260), Figure 180 (SEQ ID NO:262), Figure 182 (SEQ ID NO:264), Figure 184 (SEQ ID NO:266), Figure 186 (SEQ ID NO:268), Figure 188 (SEQ ID NO:270), Figure 190 (SEQ ID NO:272), Figure 192 (SEQ ID NO:274), Figure 194 (SEQ ID NO:276), Figure 196 (SEQ ID NO:278), Figure 198 (SEQ ID NO:281), Figure 200 (SEQ ID NO:283), Figure 202 (SEQ ID NO:285), Figure 204 (SEQ ID NO:287), Figure 206 (SEQ ID NO:289), Figure 208 (SEQ ID NO:291), Figure 210 (SEQ ID NO:293), Figure 212 (SEQ ID NO:295), Figure 214 (SEQ ID NO:297), Figure 216 (SEQ ID NO:299), Figure 218 (SEQ ID NO:301), Figure 220 (SEQ ID NO:303), Figure 226 (SEQ ID NO:309), Figure 228 (SEQ ID NO:314), Figure 230 (SEQ ID NO:319), Figure 233 (SEQ ID NO:326), Figure 235 (SEQ ID NO:334), Figure 238 (SEQ ID NO:340), Figure 240 (SEQ ID NO:345), Figure 242 (SEQ ID NO:347), Figure 244 (SEQ ID NO:349), Figure 246 (SEQ ID NO:351), Figure 248 (SEQ ID NO:353), Figure 250 (SEQ ID NO:355), Figure 252 (SEQ ID NO:357), Figure 254 (SEQ ID NO:359), Figure 256 (SEQ ID NO:361), Figure 258 (SEQ ID NO:363), Figure 260 (SEQ ID NO:365), Figure 262 (SEQ ID NO:367), Figure 264 (SEQ ID NO:369), Figure 266 (SEQ ID NO:371), Figure 268 (SEQ ID NO:373), Figure 270 (SEQ ID NO:375), Figure 272 (SEQ ID NO:377), Figure 274 (SEQ ID NO:379), Figure 276 (SEQ ID NO:381), Figure 278 (SEQ ID NO:387), Figure 280 (SEQ ID NO:389), Figure 282 (SEQ ID NO:394), Figure

284 (SEQ ID NO:399), Figure 286 (SEQ ID NO:401), Figure 288 (SEQ ID NO:403), Figure 290 (SEQ ID NO:408), Figure 292 (SEQ ID NO:410), Figure 294 (SEQ ID NO:412), Figure 296 (SEQ ID NO:414), Figure 298 (SEQ ID NO:416), Figure 300 (SEQ ID NO:418), Figure 302 (SEQ ID NO:420), Figure 304 (SEQ ID NO:422), Figure 306 (SEQ ID NO:424), Figure 308 (SEQ ID NO:495), Figure 310 (SEQ ID NO:497), Figure 312 (SEQ ID NO:499), Figure 314 (SEQ ID NO:501), Figure 316 (SEQ ID NO:503), Figure 318 (SEQ ID NO:505), Figure 320 (SEQ ID NO:507), Figure 322 (SEQ ID NO:509), Figure 324 (SEQ ID NO:511), Figure 326 (SEQ ID NO:513), Figure 328 (SEQ ID NO:515) and Figure 330 (SEQ ID NO:517).

2. The nucleic acid sequence of Claim 1, wherein said nucleotide sequence comprises a nucleotide sequence selected from the group consisting of the sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:5), Figure 5 (SEQ ID NO:7), Figure 8 (SEQ ID NO:13), Figure 11 (SEQ ID NO:19), Figure 14 (SEQ ID NO:22), Figure 17 (SEQ ID NO:27), Figure 19 (SEQ ID NO:29), Figure 22 (SEQ ID NO:32), Figure 24 (SEQ ID NO:35), Figure 26 (SEQ ID NO:40), Figure 29 (SEQ ID NO:46), Figure 31 (SEQ ID NO:51), Figure 33 (SEQ ID NO:56), Figure 35 (SEQ ID NO:61), Figure 37 (SEQ ID NO:66), Figure 40 (SEQ ID NO:72), Figure 46 (SEQ ID NO:83), Figure 48 (SEQ ID NO:94), Figure 50 (SEQ ID NO:96), Figure 52 (SEQ ID NO:98), Figure 56 (SEQ ID NO:102), Figure 63 (SEQ ID NO:112), Figure 65 (SEQ ID NO:114), Figure 67 (SEQ ID NO:116), Figure 69 (SEQ ID NO:118), Figure 71 (SEQ ID NO:123), Figure 73 (SEQ ID NO:128), Figure 75 (SEQ ID NO:134), Figure 78 (SEQ ID NO:137), Figure 82 (SEQ ID NO:145), Figure 84 (SEQ ID NO:147), Figure 87 (SEQ ID NO:150), Figure 89 (SEQ ID NO:152), Figure 92 (SEQ ID NO:155), Figure 94 (SEQ ID NO:157), Figure 96 (SEQ ID NO:159), Figure 98 (SEQ ID NO:164), Figure 100 (SEQ ID NO:166), Figure 102 (SEQ ID NO:168), Figure 104 (SEQ ID NO:170), Figure 108 (SEQ ID NO:174), Figure 110 (SEQ ID NO:176), Figure 112 (SEQ ID NO:178), Figure 114 (SEQ ID NO:180), Figure 116 (SEQ ID NO:182), Figure 119 (SEQ ID NO:188), Figure 121 (SEQ ID NO:193), Figure 124 (SEQ ID NO:196), Figure 126 (SEQ ID NO:198), Figure 128 (SEQ ID NO:200), Figure 130 (SEQ ID NO:202), Figure 132 (SEQ ID NO:204), Figure 134 (SEQ ID NO:206), Figure 136 (SEQ ID NO:208), Figure 138 (SEQ ID NO:210), Figure 140 (SEQ ID NO:212), Figure 143 (SEQ ID NO:215), Figure 146 (SEQ ID NO:218), Figure 148 (SEQ ID NO:220), Figure 150 (SEQ ID NO:222), Figure 152 (SEQ ID NO:224), Figure 154 (SEQ ID NO:226), Figure 156 (SEQ ID NO:228), Figure 158 (SEQ ID NO:230), Figure 160 (SEQ ID NO:235), Figure 162 (SEQ ID NO:240), Figure 164 (SEQ ID NO:245), Figure 166 (SEQ ID NO:247), Figure 168 (SEQ ID NO:249), Figure 170 (SEQ ID NO:252), Figure 173 (SEQ ID NO:255), Figure 175 (SEQ ID NO:257), Figure 177 (SEQ ID NO:259), Figure 179 (SEQ ID NO:261), Figure 181 (SEQ ID NO:263), Figure 183 (SEQ ID NO:265), Figure 185 (SEQ ID NO:267), Figure 187 (SEQ ID NO:269), Figure 189 (SEQ ID NO:271), Figure 191 (SEQ ID NO:273), Figure 193 (SEQ ID NO:275), Figure 195 (SEQ ID NO:277), Figure 197 (SEQ ID NO:280), Figure 199 (SEQ ID NO:282), Figure 201 (SEQ ID NO:284), Figure 203 (SEQ ID NO:286), Figure 205 (SEQ ID NO:288), Figure 207 (SEQ ID NO:290), Figure 209 (SEQ ID NO:292), Figure 211 (SEQ ID NO:294), Figure 213 (SEQ ID NO:296), Figure 215 (SEQ ID NO:298), Figure 217 (SEQ ID NO:300), Figure 219 (SEQ ID NO:302), Figure 225 (SEQ ID NO:308), Figure 227 (SEQ ID NO:313), Figure 229 (SEQ ID NO:318), Figure 232 (SEQ ID NO:325), Figure 234 (SEQ ID NO:333), Figure 237 (SEQ ID

NO:339), Figure 239 (SEQ ID NO:344), Figure 241 (SEQ ID NO:346), Figure 243 (SEQ ID NO:348), Figure 245 (SEQ ID NO:350), Figure 247 (SEQ ID NO:352), Figure 249 (SEQ ID NO:354), Figure 251 (SEQ ID NO:356), Figure 253 (SEQ ID NO:358), Figure 255 (SEQ ID NO:360), Figure 257 (SEQ ID NO:362), Figure 259 (SEQ ID NO:364), Figure 261 (SEQ ID NO:366), Figure 263 (SEQ ID NO:368), Figure 265 (SEQ ID NO:370), Figure 267 (SEQ ID NO:372), Figure 269 (SEQ ID NO:374), Figure 271 (SEQ ID NO:376), Figure 273 (SEQ ID NO:378), Figure 275 (SEQ ID NO:380), Figure 277 (SEQ ID NO:386), Figure 279 (SEQ ID NO:388), Figure 281 (SEQ ID NO:393), Figure 283 (SEQ ID NO:398), Figure 285 (SEQ ID NO:400), Figure 287 (SEQ ID NO:402), Figure 289 (SEQ ID NO:407), Figure 291 (SEQ ID NO:409), Figure 293 (SEQ ID NO:411), Figure 295 (SEQ ID NO:413), Figure 297 (SEQ ID NO:415), Figure 299 (SEQ ID NO:417), Figure 301 (SEQ ID NO:419), Figure 303 (SEQ ID NO:421), Figure 305 (SEQ ID NO:423), Figure 307 (SEQ ID NO:494), Figure 309 (SEQ ID NO:496), Figure 311 (SEQ ID NO:498), Figure 313 (SEQ ID NO:500), Figure 315 (SEQ ID NO:502), Figure 317 (SEQ ID NO:504), Figure 319 (SEQ ID NO:506), Figure 321 (SEQ ID NO:508), Figure 323 (SEQ ID NO:510), Figure 325 (SEQ ID NO:512), Figure 327 (SEQ ID NO:514) and Figure 329 (SEQ ID NO:516).

3. The nucleic acid of Claim 1, wherein said nucleotide sequence comprises a nucleotide sequence selected from the group consisting of the full-length coding sequence of the sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:5), Figure 5 (SEQ ID NO:7), Figure 8 (SEQ ID NO:13), Figure 11 (SEQ ID NO:19), Figure 14 (SEQ ID NO:22), Figure 17 (SEQ ID NO:27), Figure 19 (SEQ ID NO:29), Figure 22 (SEQ ID NO:32), Figure 24 (SEQ ID NO:35), Figure 26 (SEQ ID NO:40), Figure 29 (SEQ ID NO:46), Figure 31 (SEQ ID NO:51), Figure 33 (SEQ ID NO:56), Figure 35 (SEQ ID NO:61), Figure 37 (SEQ ID NO:66), Figure 40 (SEQ ID NO:72), Figure 46 (SEQ ID NO:83), Figure 48 (SEQ ID NO:94), Figure 50 (SEQ ID NO:96), Figure 52 (SEQ ID NO:98), Figure 56 (SEQ ID NO:102), Figure 63 (SEQ ID NO:112), Figure 65 (SEQ ID NO:114), Figure 67 (SEQ ID NO:116), Figure 69 (SEQ ID NO:118), Figure 71 (SEQ ID NO:123), Figure 73 (SEQ ID NO:128), Figure 75 (SEQ ID NO:134), Figure 78 (SEQ ID NO:137), Figure 82 (SEQ ID NO:145), Figure 84 (SEQ ID NO:147), Figure 87 (SEQ ID NO:150), Figure 89 (SEQ ID NO:152), Figure 92 (SEQ ID NO:155), Figure 94 (SEQ ID NO:157), Figure 96 (SEQ ID NO:159), Figure 98 (SEQ ID NO:164), Figure 100 (SEQ ID NO:166), Figure 102 (SEQ ID NO:168), Figure 104 (SEQ ID NO:170), Figure 108 (SEQ ID NO:174), Figure 110 (SEQ ID NO:176), Figure 112 (SEQ ID NO:178), Figure 114 (SEQ ID NO:180), Figure 116 (SEQ ID NO:182), Figure 119 (SEQ ID NO:188), Figure 121 (SEQ ID NO:193), Figure 124 (SEQ ID NO:196), Figure 126 (SEQ ID NO:198), Figure 128 (SEQ ID NO:200), Figure 130 (SEQ ID NO:202), Figure 132 (SEQ ID NO:204), Figure 134 (SEQ ID NO:206), Figure 136 (SEQ ID NO:208), Figure 138 (SEQ ID NO:210), Figure 140 (SEQ ID NO:212), Figure 143 (SEQ ID NO:215), Figure 146 (SEQ ID NO:218), Figure 148 (SEQ ID NO:220), Figure 150 (SEQ ID NO:222), Figure 152 (SEQ ID NO:224), Figure 154 (SEQ ID NO:226), Figure 156 (SEQ ID NO:228), Figure 158 (SEQ ID NO:230), Figure 160 (SEQ ID NO:235), Figure 162 (SEQ ID NO:240), Figure 164 (SEQ ID NO:245), Figure 166 (SEQ ID NO:247), Figure 168 (SEQ ID NO:249), Figure 170 (SEQ ID NO:252), Figure 173 (SEQ ID NO:255), Figure 175 (SEQ ID NO:257), Figure 177 (SEQ ID NO:259), Figure 179 (SEQ ID NO:261), Figure

181 (SEQ ID NO:263), Figure 183 (SEQ ID NO:265), Figure 185 (SEQ ID NO:267), Figure 187 (SEQ ID NO:269), Figure 189 (SEQ ID NO:271), Figure 191 (SEQ ID NO:273), Figure 193 (SEQ ID NO:275), Figure 195 (SEQ ID NO:277), Figure 197 (SEQ ID NO:280), Figure 199 (SEQ ID NO:282), Figure 201 (SEQ ID NO:284), Figure 203 (SEQ ID NO:286), Figure 205 (SEQ ID NO:288), Figure 207 (SEQ ID NO:290), Figure 209 (SEQ ID NO:292), Figure 211 (SEQ ID NO:294), Figure 213 (SEQ ID NO:296), Figure 215 (SEQ ID NO:298), Figure 217 (SEQ ID NO:300), Figure 219 (SEQ ID NO:302), Figure 225 (SEQ ID NO:308), Figure 227 (SEQ ID NO:313), Figure 229 (SEQ ID NO:318), Figure 232 (SEQ ID NO:325), Figure 234 (SEQ ID NO:333), Figure 237 (SEQ ID NO:339), Figure 239 (SEQ ID NO:344), Figure 241 (SEQ ID NO:346), Figure 243 (SEQ ID NO:348), Figure 245 (SEQ ID NO:350), Figure 247 (SEQ ID NO:352), Figure 249 (SEQ ID NO:354), Figure 251 (SEQ ID NO:356), Figure 253 (SEQ ID NO:358), Figure 255 (SEQ ID NO:360), Figure 257 (SEQ ID NO:362), Figure 259 (SEQ ID NO:364), Figure 261 (SEQ ID NO:366), Figure 263 (SEQ ID NO:368), Figure 265 (SEQ ID NO:370), Figure 267 (SEQ ID NO:372), Figure 269 (SEQ ID NO:374), Figure 271 (SEQ ID NO:376), Figure 273 (SEQ ID NO:378), Figure 275 (SEQ ID NO:380), Figure 277 (SEQ ID NO:386), Figure 279 (SEQ ID NO:388), Figure 281 (SEQ ID NO:393), Figure 283 (SEQ ID NO:398), Figure 285 (SEQ ID NO:400), Figure 287 (SEQ ID NO:402), Figure 289 (SEQ ID NO:407), Figure 291 (SEQ ID NO:409), Figure 293 (SEQ ID NO:411), Figure 295 (SEQ ID NO:413), Figure 297 (SEQ ID NO:415), Figure 299 (SEQ ID NO:417), Figure 301 (SEQ ID NO:419), Figure 303 (SEQ ID NO:421), Figure 305 (SEQ ID NO:423), Figure 307 (SEQ ID NO:494), Figure 309 (SEQ ID NO:496), Figure 311 (SEQ ID NO:498), Figure 313 (SEQ ID NO:500), Figure 315 (SEQ ID NO:502), Figure 317 (SEQ ID NO:504), Figure 319 (SEQ ID NO:506), Figure 321 (SEQ ID NO:508), Figure 323 (SEQ ID NO:510), Figure 325 (SEQ ID NO:512), Figure 327 (SEQ ID NO:514) and Figure 329 (SEQ ID NO:516).

4. Isolated nucleic acid which comprises the full-length coding sequence of the DNA deposited under any ATCC accession number shown in Table 10.

5. A vector comprising the nucleic acid of Claim 1.

6. The vector of Claim 5 operably linked to control sequences recognized by a host cell transformed with the vector.

7. A host cell comprising the vector of Claim 5.

8. The host cell of Claim 7 wherein said cell is a CHO cell.

9. The host cell of Claim 7 wherein said cell is an *E. coli*.

10. The host cell of Claim 7 wherein said cell is a yeast cell.

11. A process for producing a PRO polypeptides comprising culturing the host cell of Claim 7 under conditions suitable for expression of said PRO polypeptide and recovering said PRO polypeptide from the cell culture.

12. Isolated PRO polypeptide having at least 80% sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:6), Figure 6 (SEQ ID NO:8), Figure 9 (SEQ ID NO:14), Figure 12 (SEQ ID NO:20), Figure 15 (SEQ ID NO:23), Figure 18 (SEQ ID NO:28), Figure 20 (SEQ ID NO:30), Figure 23 (SEQ ID NO:33), Figure 25 (SEQ ID NO:36), Figure 27 (SEQ ID NO:41), Figure 30 (SEQ ID NO:47), Figure 32 (SEQ ID NO:52), Figure 34 (SEQ ID NO:57), Figure 36 (SEQ ID NO:62), Figure 38 (SEQ ID NO:67), Figure 41 (SEQ ID NO:73), Figure 47 (SEQ ID NO:84), Figure 49 (SEQ ID NO:95), Figure 51 (SEQ ID NO:97), Figure 53 (SEQ ID NO:99), Figure 57 (SEQ ID NO:103), Figure 64 (SEQ ID NO:113), Figure 66 (SEQ ID NO:115), Figure 68 (SEQ ID NO:117), Figure 70 (SEQ ID NO:119), Figure 72 (SEQ ID NO:124), Figure 74 (SEQ ID NO:129), Figure 76 (SEQ ID NO:135), Figure 79 (SEQ ID NO:138), Figure 83 (SEQ ID NO:146), Figure 85 (SEQ ID NO:148), Figure 88 (SEQ ID NO:151), Figure 90 (SEQ ID NO:153), Figure 93 (SEQ ID NO:156), Figure 95 (SEQ ID NO:158), Figure 97 (SEQ ID NO:160), Figure 99 (SEQ ID NO:165), Figure 101 (SEQ ID NO:167), Figure 103 (SEQ ID NO:169), Figure 105 (SEQ ID NO:171), Figure 109 (SEQ ID NO:175), Figure 111 (SEQ ID NO:177), Figure 113 (SEQ ID NO:179), Figure 115 (SEQ ID NO:181), Figure 117 (SEQ ID NO:183), Figure 120 (SEQ ID NO:189), Figure 122 (SEQ ID NO:194), Figure 125 (SEQ ID NO:197), Figure 127 (SEQ ID NO:199), Figure 129 (SEQ ID NO:201), Figure 131 (SEQ ID NO:203), Figure 133 (SEQ ID NO:205), Figure 135 (SEQ ID NO:207), Figure 137 (SEQ ID NO:209), Figure 139 (SEQ ID NO:211), Figure 141 (SEQ ID NO:213), Figure 144 (SEQ ID NO:216), Figure 147 (SEQ ID NO:219), Figure 149 (SEQ ID NO:221), Figure 151 (SEQ ID NO:223), Figure 153 (SEQ ID NO:225), Figure 155 (SEQ ID NO:227), Figure 157 (SEQ ID NO:229), Figure 159 (SEQ ID NO:231), Figure 161 (SEQ ID NO:236), Figure 163 (SEQ ID NO:241), Figure 165 (SEQ ID NO:246), Figure 167 (SEQ ID NO:248), Figure 169 (SEQ ID NO:250), Figure 171 (SEQ ID NO:253), Figure 174 (SEQ ID NO:256), Figure 176 (SEQ ID NO:258), Figure 178 (SEQ ID NO:260), Figure 180 (SEQ ID NO:262), Figure 182 (SEQ ID NO:264), Figure 184 (SEQ ID NO:266), Figure 186 (SEQ ID NO:268), Figure 188 (SEQ ID NO:270), Figure 190 (SEQ ID NO:272), Figure 192 (SEQ ID NO:274), Figure 194 (SEQ ID NO:276), Figure 196 (SEQ ID NO:278), Figure 198 (SEQ ID NO:281), Figure 200 (SEQ ID NO:283), Figure 202 (SEQ ID NO:285), Figure 204 (SEQ ID NO:287), Figure 206 (SEQ ID NO:289), Figure 208 (SEQ ID NO:291), Figure 210 (SEQ ID NO:293), Figure 212 (SEQ ID NO:295), Figure 214 (SEQ ID NO:297), Figure 216 (SEQ ID NO:299), Figure 218 (SEQ ID NO:301), Figure 220 (SEQ ID NO:303), Figure 226 (SEQ ID NO:309), Figure 228 (SEQ ID NO:314), Figure 230 (SEQ ID NO:319), Figure 233 (SEQ ID NO:326), Figure 235 (SEQ ID NO:334), Figure 238 (SEQ ID NO:340), Figure 240 (SEQ ID NO:345), Figure 242 (SEQ ID NO:347), Figure 244 (SEQ ID NO:349), Figure 246 (SEQ ID NO:351), Figure 248 (SEQ ID NO:353), Figure 250 (SEQ ID NO:355), Figure 252 (SEQ ID NO:357), Figure 254 (SEQ ID NO:359), Figure 256 (SEQ ID NO:361), Figure 258 (SEQ ID NO:363), Figure 260 (SEQ ID NO:365), Figure 262 (SEQ ID NO:367), Figure 264 (SEQ ID NO:369), Figure 266 (SEQ ID

NO:371), Figure 268 (SEQ ID NO:373), Figure 270 (SEQ ID NO:375), Figure 272 (SEQ ID NO:377), Figure 274 (SEQ ID NO:379), Figure 276 (SEQ ID NO:381), Figure 278 (SEQ ID NO:387), Figure 280 (SEQ ID NO:389), Figure 282 (SEQ ID NO:394), Figure 284 (SEQ ID NO:399), Figure 286 (SEQ ID NO:401), Figure 288 (SEQ ID NO:403), Figure 290 (SEQ ID NO:408), Figure 292 (SEQ ID NO:410), Figure 294 (SEQ ID NO:412), Figure 296 (SEQ ID NO:414), Figure 298 (SEQ ID NO:416), Figure 300 (SEQ ID NO:418), Figure 302 (SEQ ID NO:420), Figure 304 (SEQ ID NO:422), Figure 306 (SEQ ID NO:424), Figure 308 (SEQ ID NO:495), Figure 310 (SEQ ID NO:497), Figure 312 (SEQ ID NO:499), Figure 314 (SEQ ID NO:501), Figure 316 (SEQ ID NO:503), Figure 318 (SEQ ID NO:505), Figure 320 (SEQ ID NO:507), Figure 322 (SEQ ID NO:509), Figure 324 (SEQ ID NO:511), Figure 326 (SEQ ID NO:513), Figure 328 (SEQ ID NO:515) and Figure 330 (SEQ ID NO:517).

13. Isolated PRO polypeptide having at least 80% sequence identity to the amino acid sequence encoded by a nucleic acid molecule deposited under any ATCC accession number shown in Table 10.

14. A chimeric molecule comprising a polypeptide according to Claim 12 fused to a heterologous amino acid sequence.

15. The chimeric molecule of Claim 14 wherein said heterologous amino acid sequence is an epitope tag sequence.

16. The chimeric molecule of Claim 14 wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

17. An antibody which specifically binds to a PRO polypeptide according to Claim 12.

18. The antibody of Claim 17 wherein said antibody is a monoclonal antibody.

19. The antibody of Claim 17 wherein said antibody is a humanized antibody.

20. The antibody of Claim 17 wherein said antibody is an antibody fragment.

21. An isolated nucleic acid molecule which has at least 80% sequence identity to a nucleic acid which comprises a nucleotide sequence selected from the group consisting of that shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:5), Figure 5 (SEQ ID NO:7), Figure 8 (SEQ ID NO:13), Figure 11 (SEQ ID NO:19), Figure 14 (SEQ ID NO:22), Figure 17 (SEQ ID NO:27), Figure 19 (SEQ ID NO:29), Figure 22 (SEQ ID NO:32), Figure 24 (SEQ ID NO:35), Figure 26 (SEQ ID NO:40), Figure 29 (SEQ ID NO:46), Figure 31 (SEQ ID NO:51), Figure 33 (SEQ ID NO:56), Figure 35 (SEQ ID NO:61), Figure 37 (SEQ ID NO:66), Figure 40 (SEQ ID NO:72), Figure 46 (SEQ ID NO:83), Figure 48 (SEQ ID NO:94), Figure 50

(SEQ ID NO:96), Figure 52 (SEQ ID NO:98), Figure 56 (SEQ ID NO:102), Figure 63 (SEQ ID NO:112), Figure 65 (SEQ ID NO:114), Figure 67 (SEQ ID NO:116), Figure 69 (SEQ ID NO:118), Figure 71 (SEQ ID NO:123), Figure 73 (SEQ ID NO:128), Figure 75 (SEQ ID NO:134), Figure 78 (SEQ ID NO:137), Figure 82 (SEQ ID NO:145), Figure 84 (SEQ ID NO:147), Figure 87 (SEQ ID NO:150), Figure 89 (SEQ ID NO:152), Figure 92 (SEQ ID NO:155), Figure 94 (SEQ ID NO:157), Figure 96 (SEQ ID NO:159), Figure 98 (SEQ ID NO:164), Figure 100 (SEQ ID NO:166), Figure 102 (SEQ ID NO:168), Figure 104 (SEQ ID NO:170), Figure 108 (SEQ ID NO:174), Figure 110 (SEQ ID NO:176), Figure 112 (SEQ ID NO:178), Figure 114 (SEQ ID NO:180), Figure 116 (SEQ ID NO:182), Figure 119 (SEQ ID NO:188), Figure 121 (SEQ ID NO:193), Figure 124 (SEQ ID NO:196), Figure 126 (SEQ ID NO:198), Figure 128 (SEQ ID NO:200), Figure 130 (SEQ ID NO:202), Figure 132 (SEQ ID NO:204), Figure 134 (SEQ ID NO:206), Figure 136 (SEQ ID NO:208), Figure 138 (SEQ ID NO:210), Figure 140 (SEQ ID NO:212), Figure 143 (SEQ ID NO:215), Figure 146 (SEQ ID NO:218), Figure 148 (SEQ ID NO:220), Figure 150 (SEQ ID NO:222), Figure 152 (SEQ ID NO:224), Figure 154 (SEQ ID NO:226), Figure 156 (SEQ ID NO:228), Figure 158 (SEQ ID NO:230), Figure 160 (SEQ ID NO:235), Figure 162 (SEQ ID NO:240), Figure 164 (SEQ ID NO:245), Figure 166 (SEQ ID NO:247), Figure 168 (SEQ ID NO:249), Figure 170 (SEQ ID NO:252), Figure 173 (SEQ ID NO:255), Figure 175 (SEQ ID NO:257), Figure 177 (SEQ ID NO:259), Figure 179 (SEQ ID NO:261), Figure 181 (SEQ ID NO:263), Figure 183 (SEQ ID NO:265), Figure 185 (SEQ ID NO:267), Figure 187 (SEQ ID NO:269), Figure 189 (SEQ ID NO:271), Figure 191 (SEQ ID NO:273), Figure 193 (SEQ ID NO:275), Figure 195 (SEQ ID NO:277), Figure 197 (SEQ ID NO:280), Figure 199 (SEQ ID NO:282), Figure 201 (SEQ ID NO:284), Figure 203 (SEQ ID NO:286), Figure 205 (SEQ ID NO:288), Figure 207 (SEQ ID NO:290), Figure 209 (SEQ ID NO:292), Figure 211 (SEQ ID NO:294), Figure 213 (SEQ ID NO:296), Figure 215 (SEQ ID NO:298), Figure 217 (SEQ ID NO:300), Figure 219 (SEQ ID NO:302), Figure 225 (SEQ ID NO:308), Figure 227 (SEQ ID NO:313), Figure 229 (SEQ ID NO:318), Figure 232 (SEQ ID NO:325), Figure 234 (SEQ ID NO:333), Figure 237 (SEQ ID NO:339), Figure 239 (SEQ ID NO:344), Figure 241 (SEQ ID NO:346), Figure 243 (SEQ ID NO:348), Figure 245 (SEQ ID NO:350), Figure 247 (SEQ ID NO:352), Figure 249 (SEQ ID NO:354), Figure 251 (SEQ ID NO:356), Figure 253 (SEQ ID NO:358), Figure 255 (SEQ ID NO:360), Figure 257 (SEQ ID NO:362), Figure 259 (SEQ ID NO:364), Figure 261 (SEQ ID NO:366), Figure 263 (SEQ ID NO:368), Figure 265 (SEQ ID NO:370), Figure 267 (SEQ ID NO:372), Figure 269 (SEQ ID NO:374), Figure 271 (SEQ ID NO:376), Figure 273 (SEQ ID NO:378), Figure 275 (SEQ ID NO:380), Figure 277 (SEQ ID NO:386), Figure 279 (SEQ ID NO:388), Figure 281 (SEQ ID NO:393), Figure 283 (SEQ ID NO:398), Figure 285 (SEQ ID NO:400), Figure 287 (SEQ ID NO:402), Figure 289 (SEQ ID NO:407), Figure 291 (SEQ ID NO:409), Figure 293 (SEQ ID NO:411), Figure 295 (SEQ ID NO:413), Figure 297 (SEQ ID NO:415), Figure 299 (SEQ ID NO:417), Figure 301 (SEQ ID NO:419), Figure 303 (SEQ ID NO:421), Figure 305 (SEQ ID NO:423), Figure 307 (SEQ ID NO:494), Figure 309 (SEQ ID NO:496), Figure 311 (SEQ ID NO:498), Figure 313 (SEQ ID NO:500), Figure 315 (SEQ ID NO:502), Figure 317 (SEQ ID NO:504), Figure 319 (SEQ ID NO:506), Figure 321 (SEQ ID NO:508), Figure 323 (SEQ ID NO:510), Figure 325 (SEQ ID NO:512), Figure 327 (SEQ ID NO:514) and Figure 329 (SEQ ID NO:516).

22. An isolated nucleic acid molecule which has at least 80% sequence identity to the full-length coding sequence of a nucleotide sequence selected from the group consisting of that shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:5), Figure 5 (SEQ ID NO:7), Figure 8 (SEQ ID NO:13), Figure 11 (SEQ ID NO:19), Figure 14 (SEQ ID NO:22), Figure 17 (SEQ ID NO:27), Figure 19 (SEQ ID NO:29), Figure 22 (SEQ ID NO:32), Figure 24 (SEQ ID NO:35), Figure 26 (SEQ ID NO:40), Figure 29 (SEQ ID NO:46),
5 Figure 31 (SEQ ID NO:51), Figure 33 (SEQ ID NO:56), Figure 35 (SEQ ID NO:61), Figure 37 (SEQ ID NO:66), Figure 40 (SEQ ID NO:72), Figure 46 (SEQ ID NO:83), Figure 48 (SEQ ID NO:94), Figure 50 (SEQ ID NO:96), Figure 52 (SEQ ID NO:98), Figure 56 (SEQ ID NO:102), Figure 63 (SEQ ID NO:112), Figure 65 (SEQ ID NO:114), Figure 67 (SEQ ID NO:116), Figure 69 (SEQ ID NO:118), Figure 71 (SEQ ID NO:123), Figure 73 (SEQ ID NO:128), Figure 75 (SEQ ID NO:134), Figure 78 (SEQ ID NO:137), Figure
10 82 (SEQ ID NO:145), Figure 84 (SEQ ID NO:147), Figure 87 (SEQ ID NO:150), Figure 89 (SEQ ID NO:152), Figure 92 (SEQ ID NO:155), Figure 94 (SEQ ID NO:157), Figure 96 (SEQ ID NO:159), Figure 98 (SEQ ID NO:164), Figure 100 (SEQ ID NO:166), Figure 102 (SEQ ID NO:168), Figure 104 (SEQ ID NO:170), Figure 108 (SEQ ID NO:174), Figure 110 (SEQ ID NO:176), Figure 112 (SEQ ID NO:178), Figure 114 (SEQ ID NO:180), Figure 116 (SEQ ID NO:182), Figure 119 (SEQ ID NO:188), Figure 121 (SEQ ID NO:193), Figure 124 (SEQ ID NO:196), Figure 126 (SEQ ID NO:198), Figure 128 (SEQ ID NO:200), Figure
15 130 (SEQ ID NO:202), Figure 132 (SEQ ID NO:204), Figure 134 (SEQ ID NO:206), Figure 136 (SEQ ID NO:208), Figure 138 (SEQ ID NO:210), Figure 140 (SEQ ID NO:212), Figure 143 (SEQ ID NO:215), Figure 146 (SEQ ID NO:218), Figure 148 (SEQ ID NO:220), Figure 150 (SEQ ID NO:222), Figure 152 (SEQ ID NO:224), Figure 154 (SEQ ID NO:226), Figure 156 (SEQ ID NO:228), Figure 158 (SEQ ID NO:230), Figure
20 160 (SEQ ID NO:235), Figure 162 (SEQ ID NO:240), Figure 164 (SEQ ID NO:245), Figure 166 (SEQ ID NO:247), Figure 168 (SEQ ID NO:249), Figure 170 (SEQ ID NO:252), Figure 173 (SEQ ID NO:255), Figure 175 (SEQ ID NO:257), Figure 177 (SEQ ID NO:259), Figure 179 (SEQ ID NO:261), Figure 181 (SEQ ID NO:263), Figure 183 (SEQ ID NO:265), Figure 185 (SEQ ID NO:267), Figure 187 (SEQ ID NO:269), Figure 189 (SEQ ID NO:271), Figure 191 (SEQ ID NO:273), Figure 193 (SEQ ID NO:275), Figure 195 (SEQ ID NO:277), Figure 197 (SEQ ID NO:280), Figure 199 (SEQ ID NO:282), Figure 201 (SEQ ID NO:284), Figure
25 203 (SEQ ID NO:286), Figure 205 (SEQ ID NO:288), Figure 207 (SEQ ID NO:290), Figure 209 (SEQ ID NO:292), Figure 211 (SEQ ID NO:294), Figure 213 (SEQ ID NO:296), Figure 215 (SEQ ID NO:298), Figure 217 (SEQ ID NO:300), Figure 219 (SEQ ID NO:302), Figure 225 (SEQ ID NO:308), Figure 227 (SEQ ID NO:313), Figure 229 (SEQ ID NO:318), Figure 232 (SEQ ID NO:325), Figure 234 (SEQ ID NO:333), Figure
30 237 (SEQ ID NO:339), Figure 239 (SEQ ID NO:344), Figure 241 (SEQ ID NO:346), Figure 243 (SEQ ID NO:348), Figure 245 (SEQ ID NO:350), Figure 247 (SEQ ID NO:352), Figure 249 (SEQ ID NO:354), Figure 251 (SEQ ID NO:356), Figure 253 (SEQ ID NO:358), Figure 255 (SEQ ID NO:360), Figure 257 (SEQ ID NO:362), Figure 259 (SEQ ID NO:364), Figure 261 (SEQ ID NO:366), Figure 263 (SEQ ID NO:368), Figure 265 (SEQ ID NO:370), Figure 267 (SEQ ID NO:372), Figure 269 (SEQ ID NO:374), Figure 271 (SEQ ID NO:376), Figure 273 (SEQ ID NO:378), Figure 275 (SEQ ID NO:380), Figure 277 (SEQ ID NO:386), Figure
35 279 (SEQ ID NO:388), Figure 281 (SEQ ID NO:393), Figure 283 (SEQ ID NO:398), Figure 285 (SEQ ID NO:400), Figure 287 (SEQ ID NO:402), Figure 289 (SEQ ID NO:407), Figure 291 (SEQ ID NO:409), Figure

293 (SEQ ID NO:411), Figure 295 (SEQ ID NO:413), Figure 297 (SEQ ID NO:415), Figure 299 (SEQ ID NO:417), Figure 301 (SEQ ID NO:419), Figure 303 (SEQ ID NO:421), Figure 305 (SEQ ID NO:423), Figure 307 (SEQ ID NO:494), Figure 309 (SEQ ID NO:496), Figure 311 (SEQ ID NO:498), Figure 313 (SEQ ID NO:500), Figure 315 (SEQ ID NO:502), Figure 317 (SEQ ID NO:504), Figure 319 (SEQ ID NO:506), Figure 321 (SEQ ID NO:508), Figure 323 (SEQ ID NO:510), Figure 325 (SEQ ID NO:512), Figure 327 (SEQ ID NO:514) and Figure 329 (SEQ ID NO:516).

23. An isolated extracellular domain of of PRO polypeptide.

24. An isolated PRO polypeptide lacking its associated signal peptide.

25. An isolated polypeptide having at least about 80% amino acid sequence identity to an extracellular domain of of PRO polypeptide.

26. An isolated polypeptide having at least about 80% amino acid sequence identity to a PRO polypeptide lacking its associated signal peptide.

27. Isolated nucleic acid having at least 80% nucleic acid sequence identity to:

(a) a nucleotide sequence encoding the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:6), Figure 6 (SEQ ID NO:8), Figure 9 (SEQ ID NO:14), Figure 12 (SEQ ID NO:20), Figure 15 (SEQ ID NO:23), Figure 18 (SEQ ID NO:28), Figure 20 (SEQ ID NO:30), Figure 23 (SEQ ID NO:33), Figure 25 (SEQ ID NO:36), Figure 27 (SEQ ID NO:41), Figure 30 (SEQ ID NO:47), Figure 32 (SEQ ID NO:52), Figure 34 (SEQ ID NO:57), Figure 36 (SEQ ID NO:62), Figure 38 (SEQ ID NO:67), Figure 41 (SEQ ID NO:73), Figure 47 (SEQ ID NO:84), Figure 49 (SEQ ID NO:95), Figure 51 (SEQ ID NO:97), Figure 53 (SEQ ID NO:99), Figure 57 (SEQ ID NO:103), Figure 64 (SEQ ID NO:113), Figure 66 (SEQ ID NO:115), Figure 68 (SEQ ID NO:117), Figure 70 (SEQ ID NO:119), Figure 72 (SEQ ID NO:124), Figure 74 (SEQ ID NO:129), Figure 76 (SEQ ID NO:135), Figure 79 (SEQ ID NO:138), Figure 83 (SEQ ID NO:146), Figure 85 (SEQ ID NO:148), Figure 88 (SEQ ID NO:151), Figure 90 (SEQ ID NO:153), Figure 93 (SEQ ID NO:156), Figure 95 (SEQ ID NO:158), Figure 97 (SEQ ID NO:160), Figure 99 (SEQ ID NO:165), Figure 101 (SEQ ID NO:167), Figure 103 (SEQ ID NO:169), Figure 105 (SEQ ID NO:171), Figure 109 (SEQ ID NO:175), Figure 111 (SEQ ID NO:177), Figure 113 (SEQ ID NO:179), Figure 115 (SEQ ID NO:181), Figure 117 (SEQ ID NO:183), Figure 120 (SEQ ID NO:189), Figure 122 (SEQ ID NO:194), Figure 125 (SEQ ID NO:197), Figure 127 (SEQ ID NO:199), Figure 129 (SEQ ID NO:201), Figure 131 (SEQ ID NO:203), Figure 133 (SEQ ID NO:205), Figure 135 (SEQ ID NO:207), Figure 137 (SEQ ID NO:209), Figure 139 (SEQ ID NO:211), Figure 141 (SEQ ID NO:213), Figure 144 (SEQ ID NO:216), Figure 147 (SEQ ID NO:219), Figure 149 (SEQ ID NO:221), Figure 151 (SEQ ID NO:223), Figure 153 (SEQ ID NO:225), Figure 155 (SEQ ID NO:227), Figure 157 (SEQ ID NO:229), Figure 159 (SEQ ID NO:231), Figure 161 (SEQ ID NO:236), Figure 163 (SEQ ID NO:241), Figure 165 (SEQ ID NO:246), Figure 167 (SEQ ID NO:248), Figure

169 (SEQ ID NO:250), Figure 171 (SEQ ID NO:253), Figure 174 (SEQ ID NO:256), Figure 176 (SEQ ID NO:258), Figure 178 (SEQ ID NO:260), Figure 180 (SEQ ID NO:262), Figure 182 (SEQ ID NO:264), Figure 184 (SEQ ID NO:266), Figure 186 (SEQ ID NO:268), Figure 188 (SEQ ID NO:270), Figure 190 (SEQ ID NO:272), Figure 192 (SEQ ID NO:274), Figure 194 (SEQ ID NO:276), Figure 196 (SEQ ID NO:278), Figure 198 (SEQ ID NO:281), Figure 200 (SEQ ID NO:283), Figure 202 (SEQ ID NO:285), Figure 204 (SEQ ID NO:287), Figure 206 (SEQ ID NO:289), Figure 208 (SEQ ID NO:291), Figure 210 (SEQ ID NO:293), Figure 212 (SEQ ID NO:295), Figure 214 (SEQ ID NO:297), Figure 216 (SEQ ID NO:299), Figure 218 (SEQ ID NO:301), Figure 220 (SEQ ID NO:303), Figure 226 (SEQ ID NO:309), Figure 228 (SEQ ID NO:314), Figure 230 (SEQ ID NO:319), Figure 233 (SEQ ID NO:326), Figure 235 (SEQ ID NO:334), Figure 238 (SEQ ID NO:340), Figure 240 (SEQ ID NO:345), Figure 242 (SEQ ID NO:347), Figure 244 (SEQ ID NO:349), Figure 246 (SEQ ID NO:351), Figure 248 (SEQ ID NO:353), Figure 250 (SEQ ID NO:355), Figure 252 (SEQ ID NO:357), Figure 254 (SEQ ID NO:359), Figure 256 (SEQ ID NO:361), Figure 258 (SEQ ID NO:363), Figure 260 (SEQ ID NO:365), Figure 262 (SEQ ID NO:367), Figure 264 (SEQ ID NO:369), Figure 266 (SEQ ID NO:371), Figure 268 (SEQ ID NO:373), Figure 270 (SEQ ID NO:375), Figure 272 (SEQ ID NO:377), Figure 274 (SEQ ID NO:379), Figure 276 (SEQ ID NO:381), Figure 278 (SEQ ID NO:387), Figure 280 (SEQ ID NO:389), Figure 282 (SEQ ID NO:394), Figure 284 (SEQ ID NO:399), Figure 286 (SEQ ID NO:401), Figure 288 (SEQ ID NO:403), Figure 290 (SEQ ID NO:408), Figure 292 (SEQ ID NO:410), Figure 294 (SEQ ID NO:412), Figure 296 (SEQ ID NO:414), Figure 298 (SEQ ID NO:416), Figure 300 (SEQ ID NO:418), Figure 302 (SEQ ID NO:420), Figure 304 (SEQ ID NO:422), Figure 306 (SEQ ID NO:424), Figure 308 (SEQ ID NO:495), Figure 310 (SEQ ID NO:497), Figure 312 (SEQ ID NO:499), Figure 314 (SEQ ID NO:501), Figure 316 (SEQ ID NO:503), Figure 318 (SEQ ID NO:505), Figure 320 (SEQ ID NO:507), Figure 322 (SEQ ID NO:509), Figure 324 (SEQ ID NO:511), Figure 326 (SEQ ID NO:513), Figure 328 (SEQ ID NO:515) or Figure 330 (SEQ ID NO:517), lacking its associated signal peptide;

(b) a nucleotide sequence encoding an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:6), Figure 6 (SEQ ID NO:8), Figure 9 (SEQ ID NO:14), Figure 12 (SEQ ID NO:20), Figure 15 (SEQ ID NO:23), Figure 18 (SEQ ID NO:28), Figure 20 (SEQ ID NO:30), Figure 23 (SEQ ID NO:33), Figure 25 (SEQ ID NO:36), Figure 27 (SEQ ID NO:41), Figure 30 (SEQ ID NO:47), Figure 32 (SEQ ID NO:52), Figure 34 (SEQ ID NO:57), Figure 36 (SEQ ID NO:62), Figure 38 (SEQ ID NO:67), Figure 41 (SEQ ID NO:73), Figure 47 (SEQ ID NO:84), Figure 49 (SEQ ID NO:95), Figure 51 (SEQ ID NO:97), Figure 53 (SEQ ID NO:99), Figure 57 (SEQ ID NO:103), Figure 64 (SEQ ID NO:113), Figure 66 (SEQ ID NO:115), Figure 68 (SEQ ID NO:117), Figure 70 (SEQ ID NO:119), Figure 72 (SEQ ID NO:124), Figure 74 (SEQ ID NO:129), Figure 76 (SEQ ID NO:135), Figure 79 (SEQ ID NO:138), Figure 83 (SEQ ID NO:146), Figure 85 (SEQ ID NO:148), Figure 88 (SEQ ID NO:151), Figure 90 (SEQ ID NO:153), Figure 93 (SEQ ID NO:156), Figure 95 (SEQ ID NO:158), Figure 97 (SEQ ID NO:160), Figure 99 (SEQ ID NO:165), Figure 101 (SEQ ID NO:167), Figure 103 (SEQ ID NO:169), Figure 105 (SEQ ID NO:171), Figure 109 (SEQ ID NO:175), Figure 111 (SEQ ID NO:177), Figure 113 (SEQ ID NO:179), Figure 115 (SEQ ID NO:181), Figure 117 (SEQ ID NO:183), Figure 120 (SEQ ID NO:189), Figure 122 (SEQ ID NO:194), Figure 125 (SEQ ID NO:197), Figure 127 (SEQ ID NO:199), Figure 129 (SEQ ID

NO:201), Figure 131 (SEQ ID NO:203), Figure 133 (SEQ ID NO:205), Figure 135 (SEQ ID NO:207), Figure 137 (SEQ ID NO:209), Figure 139 (SEQ ID NO:211), Figure 141 (SEQ ID NO:213), Figure 144 (SEQ ID NO:216), Figure 147 (SEQ ID NO:219), Figure 149 (SEQ ID NO:221), Figure 151 (SEQ ID NO:223), Figure 153 (SEQ ID NO:225), Figure 155 (SEQ ID NO:227), Figure 157 (SEQ ID NO:229), Figure 159 (SEQ ID NO:231), Figure 161 (SEQ ID NO:236), Figure 163 (SEQ ID NO:241), Figure 165 (SEQ ID NO:246), Figure 167 (SEQ ID NO:248), Figure 169 (SEQ ID NO:250), Figure 171 (SEQ ID NO:253), Figure 174 (SEQ ID NO:256), Figure 176 (SEQ ID NO:258), Figure 178 (SEQ ID NO:260), Figure 180 (SEQ ID NO:262), Figure 182 (SEQ ID NO:264), Figure 184 (SEQ ID NO:266), Figure 186 (SEQ ID NO:268), Figure 188 (SEQ ID NO:270), Figure 190 (SEQ ID NO:272), Figure 192 (SEQ ID NO:274), Figure 194 (SEQ ID NO:276), Figure 196 (SEQ ID NO:278), Figure 198 (SEQ ID NO:281), Figure 200 (SEQ ID NO:283), Figure 202 (SEQ ID NO:285), Figure 204 (SEQ ID NO:287), Figure 206 (SEQ ID NO:289), Figure 208 (SEQ ID NO:291), Figure 210 (SEQ ID NO:293), Figure 212 (SEQ ID NO:295), Figure 214 (SEQ ID NO:297), Figure 216 (SEQ ID NO:299), Figure 218 (SEQ ID NO:301), Figure 220 (SEQ ID NO:303), Figure 226 (SEQ ID NO:309), Figure 228 (SEQ ID NO:314), Figure 230 (SEQ ID NO:319), Figure 233 (SEQ ID NO:326), Figure 235 (SEQ ID NO:334), Figure 238 (SEQ ID NO:340), Figure 240 (SEQ ID NO:345), Figure 242 (SEQ ID NO:347), Figure 244 (SEQ ID NO:349), Figure 246 (SEQ ID NO:351), Figure 248 (SEQ ID NO:353), Figure 250 (SEQ ID NO:355), Figure 252 (SEQ ID NO:357), Figure 254 (SEQ ID NO:359), Figure 256 (SEQ ID NO:361), Figure 258 (SEQ ID NO:363), Figure 260 (SEQ ID NO:365), Figure 262 (SEQ ID NO:367), Figure 264 (SEQ ID NO:369), Figure 266 (SEQ ID NO:371), Figure 268 (SEQ ID NO:373), Figure 270 (SEQ ID NO:375), Figure 272 (SEQ ID NO:377), Figure 274 (SEQ ID NO:379), Figure 276 (SEQ ID NO:381), Figure 278 (SEQ ID NO:387), Figure 280 (SEQ ID NO:389), Figure 282 (SEQ ID NO:394), Figure 284 (SEQ ID NO:399), Figure 286 (SEQ ID NO:401), Figure 288 (SEQ ID NO:403), Figure 290 (SEQ ID NO:408), Figure 292 (SEQ ID NO:410), Figure 294 (SEQ ID NO:412), Figure 296 (SEQ ID NO:414), Figure 298 (SEQ ID NO:416), Figure 300 (SEQ ID NO:418), Figure 302 (SEQ ID NO:420), Figure 304 (SEQ ID NO:422), Figure 306 (SEQ ID NO:424), Figure 308 (SEQ ID NO:495), Figure 310 (SEQ ID NO:497), Figure 312 (SEQ ID NO:499), Figure 314 (SEQ ID NO:501), Figure 316 (SEQ ID NO:503), Figure 318 (SEQ ID NO:505), Figure 320 (SEQ ID NO:507), Figure 322 (SEQ ID NO:509), Figure 324 (SEQ ID NO:511), Figure 326 (SEQ ID NO:513), Figure 328 (SEQ ID NO:515) or Figure 330 (SEQ ID NO:517), with its associated signal peptide; or

(c) a nucleotide sequence encoding an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:6), Figure 6 (SEQ ID NO:8), Figure 9 (SEQ ID NO:14), Figure 12 (SEQ ID NO:20), Figure 15 (SEQ ID NO:23), Figure 18 (SEQ ID NO:28), Figure 20 (SEQ ID NO:30), Figure 23 (SEQ ID NO:33), Figure 25 (SEQ ID NO:36), Figure 27 (SEQ ID NO:41), Figure 30 (SEQ ID NO:47), Figure 32 (SEQ ID NO:52), Figure 34 (SEQ ID NO:57), Figure 36 (SEQ ID NO:62), Figure 38 (SEQ ID NO:67), Figure 41 (SEQ ID NO:73), Figure 47 (SEQ ID NO:84), Figure 49 (SEQ ID NO:95), Figure 51 (SEQ ID NO:97), Figure 53 (SEQ ID NO:99), Figure 57 (SEQ ID NO:103), Figure 64 (SEQ ID NO:113), Figure 66 (SEQ ID NO:115), Figure 68 (SEQ ID NO:117), Figure 70 (SEQ ID NO:119), Figure 72 (SEQ ID NO:124), Figure 74 (SEQ ID NO:129), Figure 76 (SEQ ID NO:135), Figure 79 (SEQ ID NO:138), Figure 83 (SEQ ID NO:146), Figure 85 (SEQ ID NO:148), Figure 88 (SEQ ID NO:151), Figure

90 (SEQ ID NO:153), Figure 93 (SEQ ID NO:156), Figure 95 (SEQ ID NO:158), Figure 97 (SEQ ID NO:160), Figure 99 (SEQ ID NO:165), Figure 101 (SEQ ID NO:167), Figure 103 (SEQ ID NO:169), Figure 105 (SEQ ID NO:171), Figure 109 (SEQ ID NO:175), Figure 111 (SEQ ID NO:177), Figure 113 (SEQ ID NO:179), Figure 115 (SEQ ID NO:181), Figure 117 (SEQ ID NO:183), Figure 120 (SEQ ID NO:189), Figure 122 (SEQ ID NO:194), Figure 125 (SEQ ID NO:197), Figure 127 (SEQ ID NO:199), Figure 129 (SEQ ID NO:201), Figure 131 (SEQ ID NO:203), Figure 133 (SEQ ID NO:205), Figure 135 (SEQ ID NO:207), Figure 137 (SEQ ID NO:209), Figure 139 (SEQ ID NO:211), Figure 141 (SEQ ID NO:213), Figure 144 (SEQ ID NO:216), Figure 147 (SEQ ID NO:219), Figure 149 (SEQ ID NO:221), Figure 151 (SEQ ID NO:223), Figure 153 (SEQ ID NO:225), Figure 155 (SEQ ID NO:227), Figure 157 (SEQ ID NO:229), Figure 159 (SEQ ID NO:231), Figure 161 (SEQ ID NO:236), Figure 163 (SEQ ID NO:241), Figure 165 (SEQ ID NO:246), Figure 167 (SEQ ID NO:248), Figure 169 (SEQ ID NO:250), Figure 171 (SEQ ID NO:253), Figure 174 (SEQ ID NO:256), Figure 176 (SEQ ID NO:258), Figure 178 (SEQ ID NO:260), Figure 180 (SEQ ID NO:262), Figure 182 (SEQ ID NO:264), Figure 184 (SEQ ID NO:266), Figure 186 (SEQ ID NO:268), Figure 188 (SEQ ID NO:270), Figure 190 (SEQ ID NO:272), Figure 192 (SEQ ID NO:274), Figure 194 (SEQ ID NO:276), Figure 196 (SEQ ID NO:278), Figure 198 (SEQ ID NO:281), Figure 200 (SEQ ID NO:283), Figure 202 (SEQ ID NO:285), Figure 204 (SEQ ID NO:287), Figure 206 (SEQ ID NO:289), Figure 208 (SEQ ID NO:291), Figure 210 (SEQ ID NO:293), Figure 212 (SEQ ID NO:295), Figure 214 (SEQ ID NO:297), Figure 216 (SEQ ID NO:299), Figure 218 (SEQ ID NO:301), Figure 220 (SEQ ID NO:303), Figure 226 (SEQ ID NO:309), Figure 228 (SEQ ID NO:314), Figure 230 (SEQ ID NO:319), Figure 233 (SEQ ID NO:326), Figure 235 (SEQ ID NO:334), Figure 238 (SEQ ID NO:340), Figure 240 (SEQ ID NO:345), Figure 242 (SEQ ID NO:347), Figure 244 (SEQ ID NO:349), Figure 246 (SEQ ID NO:351), Figure 248 (SEQ ID NO:353), Figure 250 (SEQ ID NO:355), Figure 252 (SEQ ID NO:357), Figure 254 (SEQ ID NO:359), Figure 256 (SEQ ID NO:361), Figure 258 (SEQ ID NO:363), Figure 260 (SEQ ID NO:365), Figure 262 (SEQ ID NO:367), Figure 264 (SEQ ID NO:369), Figure 266 (SEQ ID NO:371), Figure 268 (SEQ ID NO:373), Figure 270 (SEQ ID NO:375), Figure 272 (SEQ ID NO:377), Figure 274 (SEQ ID NO:379), Figure 276 (SEQ ID NO:381), Figure 278 (SEQ ID NO:387), Figure 280 (SEQ ID NO:389), Figure 282 (SEQ ID NO:394), Figure 284 (SEQ ID NO:399), Figure 286 (SEQ ID NO:401), Figure 288 (SEQ ID NO:403), Figure 290 (SEQ ID NO:408), Figure 292 (SEQ ID NO:410), Figure 294 (SEQ ID NO:412), Figure 296 (SEQ ID NO:414), Figure 298 (SEQ ID NO:416), Figure 300 (SEQ ID NO:418), Figure 302 (SEQ ID NO:420), Figure 304 (SEQ ID NO:422), Figure 306 (SEQ ID NO:424), Figure 308 (SEQ ID NO:495), Figure 310 (SEQ ID NO:497), Figure 312 (SEQ ID NO:499), Figure 314 (SEQ ID NO:501), Figure 316 (SEQ ID NO:503), Figure 318 (SEQ ID NO:505), Figure 320 (SEQ ID NO:507), Figure 322 (SEQ ID NO:509), Figure 324 (SEQ ID NO:511), Figure 326 (SEQ ID NO:513), Figure 328 (SEQ ID NO:515) or Figure 330 (SEQ ID NO:517), lacking its associated signal peptide.

28. An isolated polypeptide having at least 80% amino acid sequence identity to:

(a) the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:6), Figure 6 (SEQ ID NO:8), Figure 9 (SEQ ID NO:14), Figure 12 (SEQ ID NO:20), Figure 15 (SEQ ID NO:23), Figure 18 (SEQ ID NO:28), Figure 20 (SEQ ID NO:30), Figure 23 (SEQ ID NO:33), Figure 25 (SEQ ID NO:36),

Figure 27 (SEQ ID NO:41), Figure 30 (SEQ ID NO:47), Figure 32 (SEQ ID NO:52), Figure 34 (SEQ ID NO:57), Figure 36 (SEQ ID NO:62), Figure 38 (SEQ ID NO:67), Figure 41 (SEQ ID NO:73), Figure 47 (SEQ ID NO:84), Figure 49 (SEQ ID NO:95), Figure 51 (SEQ ID NO:97), Figure 53 (SEQ ID NO:99), Figure 57 (SEQ ID NO:103), Figure 64 (SEQ ID NO:113), Figure 66 (SEQ ID NO:115), Figure 68 (SEQ ID NO:117), Figure 70 (SEQ ID NO:119), Figure 72 (SEQ ID NO:124), Figure 74 (SEQ ID NO:129), Figure 76 (SEQ ID NO:135), Figure 79 (SEQ ID NO:138), Figure 83 (SEQ ID NO:146), Figure 85 (SEQ ID NO:148), Figure 88 (SEQ ID NO:151), Figure 90 (SEQ ID NO:153), Figure 93 (SEQ ID NO:156), Figure 95 (SEQ ID NO:158), Figure 97 (SEQ ID NO:160), Figure 99 (SEQ ID NO:165), Figure 101 (SEQ ID NO:167), Figure 103 (SEQ ID NO:169), Figure 105 (SEQ ID NO:171), Figure 109 (SEQ ID NO:175), Figure 111 (SEQ ID NO:177), Figure 113 (SEQ ID NO:179), Figure 115 (SEQ ID NO:181), Figure 117 (SEQ ID NO:183), Figure 120 (SEQ ID NO:189), Figure 122 (SEQ ID NO:194), Figure 125 (SEQ ID NO:197), Figure 127 (SEQ ID NO:199), Figure 129 (SEQ ID NO:201), Figure 131 (SEQ ID NO:203), Figure 133 (SEQ ID NO:205), Figure 135 (SEQ ID NO:207), Figure 137 (SEQ ID NO:209), Figure 139 (SEQ ID NO:211), Figure 141 (SEQ ID NO:213), Figure 144 (SEQ ID NO:216), Figure 147 (SEQ ID NO:219), Figure 149 (SEQ ID NO:221), Figure 151 (SEQ ID NO:223), Figure 153 (SEQ ID NO:225), Figure 155 (SEQ ID NO:227), Figure 157 (SEQ ID NO:229), Figure 159 (SEQ ID NO:231), Figure 161 (SEQ ID NO:236), Figure 163 (SEQ ID NO:241), Figure 165 (SEQ ID NO:246), Figure 167 (SEQ ID NO:248), Figure 169 (SEQ ID NO:250), Figure 171 (SEQ ID NO:253), Figure 174 (SEQ ID NO:256), Figure 176 (SEQ ID NO:258), Figure 178 (SEQ ID NO:260), Figure 180 (SEQ ID NO:262), Figure 182 (SEQ ID NO:264), Figure 184 (SEQ ID NO:266), Figure 186 (SEQ ID NO:268), Figure 188 (SEQ ID NO:270), Figure 190 (SEQ ID NO:272), Figure 192 (SEQ ID NO:274), Figure 194 (SEQ ID NO:276), Figure 196 (SEQ ID NO:278), Figure 198 (SEQ ID NO:281), Figure 200 (SEQ ID NO:283), Figure 202 (SEQ ID NO:285), Figure 204 (SEQ ID NO:287), Figure 206 (SEQ ID NO:289), Figure 208 (SEQ ID NO:291), Figure 210 (SEQ ID NO:293), Figure 212 (SEQ ID NO:295), Figure 214 (SEQ ID NO:297), Figure 216 (SEQ ID NO:299), Figure 218 (SEQ ID NO:301), Figure 220 (SEQ ID NO:303), Figure 226 (SEQ ID NO:309), Figure 228 (SEQ ID NO:314), Figure 230 (SEQ ID NO:319), Figure 233 (SEQ ID NO:326), Figure 235 (SEQ ID NO:334), Figure 238 (SEQ ID NO:340), Figure 240 (SEQ ID NO:345), Figure 242 (SEQ ID NO:347), Figure 244 (SEQ ID NO:349), Figure 246 (SEQ ID NO:351), Figure 248 (SEQ ID NO:353), Figure 250 (SEQ ID NO:355), Figure 252 (SEQ ID NO:357), Figure 254 (SEQ ID NO:359), Figure 256 (SEQ ID NO:361), Figure 258 (SEQ ID NO:363), Figure 260 (SEQ ID NO:365), Figure 262 (SEQ ID NO:367), Figure 264 (SEQ ID NO:369), Figure 266 (SEQ ID NO:371), Figure 268 (SEQ ID NO:373), Figure 270 (SEQ ID NO:375), Figure 272 (SEQ ID NO:377), Figure 274 (SEQ ID NO:379), Figure 276 (SEQ ID NO:381), Figure 278 (SEQ ID NO:387), Figure 280 (SEQ ID NO:389), Figure 282 (SEQ ID NO:394), Figure 284 (SEQ ID NO:399), Figure 286 (SEQ ID NO:401), Figure 288 (SEQ ID NO:403), Figure 290 (SEQ ID NO:408), Figure 292 (SEQ ID NO:410), Figure 294 (SEQ ID NO:412), Figure 296 (SEQ ID NO:414), Figure 298 (SEQ ID NO:416), Figure 300 (SEQ ID NO:418), Figure 302 (SEQ ID NO:420), Figure 304 (SEQ ID NO:422), Figure 306 (SEQ ID NO:424), Figure 308 (SEQ ID NO:495), Figure 310 (SEQ ID NO:497), Figure 312 (SEQ ID NO:499), Figure 314 (SEQ ID NO:501), Figure 316 (SEQ ID NO:503), Figure 318 (SEQ ID NO:505), Figure 320 (SEQ ID NO:507), Figure 322 (SEQ ID NO:509), Figure 324 (SEQ ID

NO:511), Figure 326 (SEQ ID NO:513), Figure 328 (SEQ ID NO:515) or Figure 330 (SEQ ID NO:517), lacking its associated signal peptide;

(b) an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:6), Figure 6 (SEQ ID NO:8), Figure 9 (SEQ ID NO:14), Figure 12 (SEQ ID NO:20), Figure 15 (SEQ ID NO:23), Figure 18 (SEQ ID NO:28), Figure 20 (SEQ ID NO:30), Figure 23 (SEQ ID NO:33), Figure 25 (SEQ ID NO:36), Figure 27 (SEQ ID NO:41), Figure 30 (SEQ ID NO:47), Figure 32 (SEQ ID NO:52), Figure 34 (SEQ ID NO:57), Figure 36 (SEQ ID NO:62), Figure 38 (SEQ ID NO:67), Figure 41 (SEQ ID NO:73), Figure 47 (SEQ ID NO:84), Figure 49 (SEQ ID NO:95), Figure 51 (SEQ ID NO:97), Figure 53 (SEQ ID NO:99), Figure 57 (SEQ ID NO:103), Figure 64 (SEQ ID NO:113), Figure 66 (SEQ ID NO:115), Figure 68 (SEQ ID NO:117), Figure 70 (SEQ ID NO:119), Figure 72 (SEQ ID NO:124), Figure 74 (SEQ ID NO:129), Figure 76 (SEQ ID NO:135), Figure 79 (SEQ ID NO:138), Figure 83 (SEQ ID NO:146), Figure 85 (SEQ ID NO:148), Figure 88 (SEQ ID NO:151), Figure 90 (SEQ ID NO:153), Figure 93 (SEQ ID NO:156), Figure 95 (SEQ ID NO:158), Figure 97 (SEQ ID NO:160), Figure 99 (SEQ ID NO:165), Figure 101 (SEQ ID NO:167), Figure 103 (SEQ ID NO:169), Figure 105 (SEQ ID NO:171), Figure 109 (SEQ ID NO:175), Figure 111 (SEQ ID NO:177), Figure 113 (SEQ ID NO:179), Figure 115 (SEQ ID NO:181), Figure 117 (SEQ ID NO:183), Figure 120 (SEQ ID NO:189), Figure 122 (SEQ ID NO:194), Figure 125 (SEQ ID NO:197), Figure 127 (SEQ ID NO:199), Figure 129 (SEQ ID NO:201), Figure 131 (SEQ ID NO:203), Figure 133 (SEQ ID NO:205), Figure 135 (SEQ ID NO:207), Figure 137 (SEQ ID NO:209), Figure 139 (SEQ ID NO:211), Figure 141 (SEQ ID NO:213), Figure 144 (SEQ ID NO:216), Figure 147 (SEQ ID NO:219), Figure 149 (SEQ ID NO:221), Figure 151 (SEQ ID NO:223), Figure 153 (SEQ ID NO:225), Figure 155 (SEQ ID NO:227), Figure 157 (SEQ ID NO:229), Figure 159 (SEQ ID NO:231), Figure 161 (SEQ ID NO:236), Figure 163 (SEQ ID NO:241), Figure 165 (SEQ ID NO:246), Figure 167 (SEQ ID NO:248), Figure 169 (SEQ ID NO:250), Figure 171 (SEQ ID NO:253), Figure 174 (SEQ ID NO:256), Figure 176 (SEQ ID NO:258), Figure 178 (SEQ ID NO:260), Figure 180 (SEQ ID NO:262), Figure 182 (SEQ ID NO:264), Figure 184 (SEQ ID NO:266), Figure 186 (SEQ ID NO:268), Figure 188 (SEQ ID NO:270), Figure 190 (SEQ ID NO:272), Figure 192 (SEQ ID NO:274), Figure 194 (SEQ ID NO:276), Figure 196 (SEQ ID NO:278), Figure 198 (SEQ ID NO:281), Figure 200 (SEQ ID NO:283), Figure 202 (SEQ ID NO:285), Figure 204 (SEQ ID NO:287), Figure 206 (SEQ ID NO:289), Figure 208 (SEQ ID NO:291), Figure 210 (SEQ ID NO:293), Figure 212 (SEQ ID NO:295), Figure 214 (SEQ ID NO:297), Figure 216 (SEQ ID NO:299), Figure 218 (SEQ ID NO:301), Figure 220 (SEQ ID NO:303), Figure 226 (SEQ ID NO:309), Figure 228 (SEQ ID NO:314), Figure 230 (SEQ ID NO:319), Figure 233 (SEQ ID NO:326), Figure 235 (SEQ ID NO:334), Figure 238 (SEQ ID NO:340), Figure 240 (SEQ ID NO:345), Figure 242 (SEQ ID NO:347), Figure 244 (SEQ ID NO:349), Figure 246 (SEQ ID NO:351), Figure 248 (SEQ ID NO:353), Figure 250 (SEQ ID NO:355), Figure 252 (SEQ ID NO:357), Figure 254 (SEQ ID NO:359), Figure 256 (SEQ ID NO:361), Figure 258 (SEQ ID NO:363), Figure 260 (SEQ ID NO:365), Figure 262 (SEQ ID NO:367), Figure 264 (SEQ ID NO:369), Figure 266 (SEQ ID NO:371), Figure 268 (SEQ ID NO:373), Figure 270 (SEQ ID NO:375), Figure 272 (SEQ ID NO:377), Figure 274 (SEQ ID NO:379), Figure 276 (SEQ ID NO:381), Figure 278 (SEQ ID NO:387), Figure 280 (SEQ ID NO:389), Figure 282 (SEQ ID NO:394), Figure 284 (SEQ ID NO:399), Figure 286 (SEQ ID NO:401), Figure 288 (SEQ ID

NO:403), Figure 290 (SEQ ID NO:408), Figure 292 (SEQ ID NO:410), Figure 294 (SEQ ID NO:412), Figure 296 (SEQ ID NO:414), Figure 298 (SEQ ID NO:416), Figure 300 (SEQ ID NO:418), Figure 302 (SEQ ID NO:420), Figure 304 (SEQ ID NO:422), Figure 306 (SEQ ID NO:424), Figure 308 (SEQ ID NO:495), Figure 310 (SEQ ID NO:497), Figure 312 (SEQ ID NO:499), Figure 314 (SEQ ID NO:501), Figure 316 (SEQ ID NO:503), Figure 318 (SEQ ID NO:505), Figure 320 (SEQ ID NO:507), Figure 322 (SEQ ID NO:509), Figure 324 (SEQ ID NO:511), Figure 326 (SEQ ID NO:513), Figure 328 (SEQ ID NO:515) or Figure 330 (SEQ ID NO:517), with its associated signal peptide; or

(c) an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:6), Figure 6 (SEQ ID NO:8), Figure 9 (SEQ ID NO:14), Figure 12 (SEQ ID NO:20), Figure 15 (SEQ ID NO:23), Figure 18 (SEQ ID NO:28), Figure 20 (SEQ ID NO:30), Figure 23 (SEQ ID NO:33), Figure 25 (SEQ ID NO:36), Figure 27 (SEQ ID NO:41), Figure 30 (SEQ ID NO:47), Figure 32 (SEQ ID NO:52), Figure 34 (SEQ ID NO:57), Figure 36 (SEQ ID NO:62), Figure 38 (SEQ ID NO:67), Figure 41 (SEQ ID NO:73), Figure 47 (SEQ ID NO:84), Figure 49 (SEQ ID NO:95), Figure 51 (SEQ ID NO:97), Figure 53 (SEQ ID NO:99), Figure 57 (SEQ ID NO:103), Figure 64 (SEQ ID NO:113), Figure 66 (SEQ ID NO:115), Figure 68 (SEQ ID NO:117), Figure 70 (SEQ ID NO:119), Figure 72 (SEQ ID NO:124), Figure 74 (SEQ ID NO:129), Figure 76 (SEQ ID NO:135), Figure 79 (SEQ ID NO:138), Figure 83 (SEQ ID NO:146), Figure 85 (SEQ ID NO:148), Figure 88 (SEQ ID NO:151), Figure 90 (SEQ ID NO:153), Figure 93 (SEQ ID NO:156), Figure 95 (SEQ ID NO:158), Figure 97 (SEQ ID NO:160), Figure 99 (SEQ ID NO:165), Figure 101 (SEQ ID NO:167), Figure 103 (SEQ ID NO:169), Figure 105 (SEQ ID NO:171), Figure 109 (SEQ ID NO:175), Figure 111 (SEQ ID NO:177), Figure 113 (SEQ ID NO:179), Figure 115 (SEQ ID NO:181), Figure 117 (SEQ ID NO:183), Figure 120 (SEQ ID NO:189), Figure 122 (SEQ ID NO:194), Figure 125 (SEQ ID NO:197), Figure 127 (SEQ ID NO:199), Figure 129 (SEQ ID NO:201), Figure 131 (SEQ ID NO:203), Figure 133 (SEQ ID NO:205), Figure 135 (SEQ ID NO:207), Figure 137 (SEQ ID NO:209), Figure 139 (SEQ ID NO:211), Figure 141 (SEQ ID NO:213), Figure 144 (SEQ ID NO:216), Figure 147 (SEQ ID NO:219), Figure 149 (SEQ ID NO:221), Figure 151 (SEQ ID NO:223), Figure 153 (SEQ ID NO:225), Figure 155 (SEQ ID NO:227), Figure 157 (SEQ ID NO:229), Figure 159 (SEQ ID NO:231), Figure 161 (SEQ ID NO:236), Figure 163 (SEQ ID NO:241), Figure 165 (SEQ ID NO:246), Figure 167 (SEQ ID NO:248), Figure 169 (SEQ ID NO:250), Figure 171 (SEQ ID NO:253), Figure 174 (SEQ ID NO:256), Figure 176 (SEQ ID NO:258), Figure 178 (SEQ ID NO:260), Figure 180 (SEQ ID NO:262), Figure 182 (SEQ ID NO:264), Figure 184 (SEQ ID NO:266), Figure 186 (SEQ ID NO:268), Figure 188 (SEQ ID NO:270), Figure 190 (SEQ ID NO:272), Figure 192 (SEQ ID NO:274), Figure 194 (SEQ ID NO:276), Figure 196 (SEQ ID NO:278), Figure 198 (SEQ ID NO:281), Figure 200 (SEQ ID NO:283), Figure 202 (SEQ ID NO:285), Figure 204 (SEQ ID NO:287), Figure 206 (SEQ ID NO:289), Figure 208 (SEQ ID NO:291), Figure 210 (SEQ ID NO:293), Figure 212 (SEQ ID NO:295), Figure 214 (SEQ ID NO:297), Figure 216 (SEQ ID NO:299), Figure 218 (SEQ ID NO:301), Figure 220 (SEQ ID NO:303), Figure 226 (SEQ ID NO:309), Figure 228 (SEQ ID NO:314), Figure 230 (SEQ ID NO:319), Figure 233 (SEQ ID NO:326), Figure 235 (SEQ ID NO:334), Figure 238 (SEQ ID NO:340), Figure 240 (SEQ ID NO:345), Figure 242 (SEQ ID NO:347), Figure 244 (SEQ ID NO:349), Figure 246 (SEQ ID NO:351), Figure 248 (SEQ ID NO:353), Figure 250 (SEQ ID NO:355), Figure 252 (SEQ ID NO:357), Figure

254 (SEQ ID NO:359), Figure 256 (SEQ ID NO:361), Figure 258 (SEQ ID NO:363), Figure 260 (SEQ ID NO:365), Figure 262 (SEQ ID NO:367), Figure 264 (SEQ ID NO:369), Figure 266 (SEQ ID NO:371), Figure 268 (SEQ ID NO:373), Figure 270 (SEQ ID NO:375), Figure 272 (SEQ ID NO:377), Figure 274 (SEQ ID NO:379), Figure 276 (SEQ ID NO:381), Figure 278 (SEQ ID NO:387), Figure 280 (SEQ ID NO:389), Figure 282 (SEQ ID NO:394), Figure 284 (SEQ ID NO:399), Figure 286 (SEQ ID NO:401), Figure 288 (SEQ ID NO:403), Figure 290 (SEQ ID NO:408), Figure 292 (SEQ ID NO:410), Figure 294 (SEQ ID NO:412), Figure 296 (SEQ ID NO:414), Figure 298 (SEQ ID NO:416), Figure 300 (SEQ ID NO:418), Figure 302 (SEQ ID NO:420), Figure 304 (SEQ ID NO:422), Figure 306 (SEQ ID NO:424), Figure 308 (SEQ ID NO:495), Figure 310 (SEQ ID NO:497), Figure 312 (SEQ ID NO:499), Figure 314 (SEQ ID NO:501), Figure 316 (SEQ ID NO:503), Figure 318 (SEQ ID NO:505), Figure 320 (SEQ ID NO:507), Figure 322 (SEQ ID NO:509), Figure 324 (SEQ ID NO:511), Figure 326 (SEQ ID NO:513), Figure 328 (SEQ ID NO:515) or Figure 330 (SEQ ID NO:517), lacking its associated signal peptide.

29. A method of detecting a PRO943 polypeptide in a sample suspected of containing a PRO943 polypeptide, said method comprising contacting said sample with a PRO183, PRO184 or PRO185 polypeptide and determining the formation of a PRO943/PRO183, PRO184 or PRO185 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO943 polypeptide in said sample.

30. The method according to Claim 29, wherein said sample comprises cells suspected of expressing said PRO943 polypeptide.

31. The method according to Claim 29, wherein said PRO183, PRO184 or PRO185 polypeptide is labeled with a detectable label.

32. The method according to Claim 29, wherein said PRO183, PRO184 or PRO185 polypeptide is attached to a solid support.

33. A method of detecting a PRO183, PRO184 or PRO185 polypeptide in a sample suspected of containing a PRO183, PRO184 or PRO185 polypeptide, said method comprising contacting said sample with a PRO943 polypeptide and determining the formation of a PRO943/PRO183, PRO184 or PRO185 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO183, PRO184 or PRO185 polypeptide in said sample.

34. The method according to Claim 33, wherein said sample comprises cells suspected of expressing said PRO183, PRO184 or PRO185 polypeptide.

35. The method according to Claim 33, wherein said PRO943 polypeptide is labeled with a detectable label.

36. The method according to Claim 33, wherein said PRO943 polypeptide is attached to a solid support.

37. A method of detecting a PRO331 polypeptide in a sample suspected of containing a PRO331 polypeptide, said method comprising contacting said sample with a PRO1133 polypeptide and determining the formation of a PRO331/PRO1133 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO331 polypeptide in said sample.

38. The method according to Claim 37, wherein said sample comprises cells suspected of expressing said PRO331 polypeptide.

39. The method according to Claim 37, wherein said PRO1133 polypeptide is labeled with a detectable label.

40. The method according to Claim 37, wherein said PRO1133 polypeptide is attached to a solid support.

41. A method of detecting a PRO1133 polypeptide in a sample suspected of containing a PRO1133 polypeptide, said method comprising contacting said sample with a PRO331 polypeptide and determining the formation of a PRO331/PRO1133 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO1133 polypeptide in said sample.

42. The method according to Claim 41, wherein said sample comprises cells suspected of expressing said PRO1133 polypeptide.

43. The method according to Claim 41, wherein said PRO331 polypeptide is labeled with a detectable label.

44. The method according to Claim 41, wherein said PRO331 polypeptide is attached to a solid support.

45. A method of detecting a PRO363 or PRO5723 polypeptide in a sample suspected of containing a PRO363 or PRO5723 polypeptide, said method comprising contacting said sample with a PRO1387 polypeptide and determining the formation of a PRO363 or PRO5723/PRO1387 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO363

or PRO5723 polypeptide in said sample.

46. The method according to Claim 45, wherein said sample comprises cells suspected of expressing said PRO363 or PRO5723 polypeptide.

5 47. The method according to Claim 45, wherein said PRO1387 polypeptide is labeled with a detectable label.

48. The method according to Claim 45, wherein said PRO1387 polypeptide is attached to a solid support.

10 49. A method of detecting a PRO1387 polypeptide in a sample suspected of containing a PRO1387 polypeptide, said method comprising contacting said sample with a PRO363 or PRO5723 polypeptide and determining the formation of a PRO363 or PRO5723/PRO1387 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO1387 polypeptide in said sample.

15 50. The method according to Claim 49, wherein said sample comprises cells suspected of expressing said PRO1387 polypeptide.

20 51. The method according to Claim 49, wherein said PRO363 or PRO5723 polypeptide is labeled with a detectable label.

52. The method according to Claim 49, wherein said PRO363 or PRO5723 polypeptide is attached to a solid support.

25 53. A method of detecting a PRO1114 polypeptide in a sample suspected of containing a PRO1114 polypeptide, said method comprising contacting said sample with a PRO3301 or PRO9940 polypeptide and determining the formation of a PRO1114/PRO3301 or PRO9940 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO1114 polypeptide in said sample.

30 54. The method according to Claim 53, wherein said sample comprises cells suspected of expressing said PRO1114 polypeptide.

35 55. The method according to Claim 53, wherein said PRO3301 or PRO9940 polypeptide is labeled with a detectable label.

56. The method according to Claim 53, wherein said PRO3301 or PRO9940 polypeptide is attached to a solid support.

57. A method of detecting a PRO3301 or PRO9940 polypeptide in a sample suspected of containing a PRO3301 or PRO9940 polypeptide, said method comprising contacting said sample with a PRO1114 polypeptide and determining the formation of a PRO3301 or PRO9940/PRO1114 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO3301 or PRO9940 polypeptide in said sample.

58. The method according to Claim 57, wherein said sample comprises cells suspected of expressing said PRO3301 or PRO9940 polypeptide.

59. The method according to Claim 57, wherein said PRO1114 polypeptide is labeled with a detectable label.

60. The method according to Claim 57, wherein said PRO1114 polypeptide is attached to a solid support.

61. A method of detecting a PRO1181 polypeptide in a sample suspected of containing a PRO1181 polypeptide, said method comprising contacting said sample with a PRO7170, PRO361 or PRO846 polypeptide and determining the formation of a PRO1181/PRO7170, PRO361 or PRO846 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO1181 polypeptide in said sample.

62. The method according to Claim 61, wherein said sample comprises cells suspected of expressing said PRO1181 polypeptide.

63. The method according to Claim 61, wherein said PRO7170, PRO361 or PRO846 polypeptide is labeled with a detectable label.

64. The method according to Claim 61, wherein said PRO7170, PRO361 or PRO846 polypeptide is attached to a solid support.

65. A method of detecting a PRO7170, PRO361 or PRO846 polypeptide in a sample suspected of containing a PRO7170, PRO361 or PRO846 polypeptide, said method comprising contacting said sample with a PRO1181 polypeptide and determining the formation of a PRO1181/PRO7170, PRO361 or PRO846 polypeptide conjugate in said sample, wherein the formation of said conjugate is indicative of the presence of a PRO7170, PRO361 or PRO846 polypeptide in said sample.

66. The method according to Claim 65, wherein said sample comprises cells suspected of expressing said PRO7170, PRO361 or PRO846 polypeptide.

67. The method according to Claim 65, wherein said PRO1181 polypeptide is labeled with a detectable label.

68. The method according to Claim 65, wherein said PRO1181 polypeptide is attached to a solid support.

69. A method of linking a bioactive molecule to a cell expressing a PRO943 polypeptide, said method comprising contacting said cell with a PRO183, PRO184 or PRO185 polypeptide that is bound to said bioactive molecule and allowing said PRO943 and PRO183, PRO184 or PRO185 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

70. The method according to Claim 69, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

71. The method according to Claim 69, wherein said bioactive molecule causes the death of said cell.

72. A method of linking a bioactive molecule to a cell expressing a PRO183, PRO184 or PRO185 polypeptide, said method comprising contacting said cell with a PRO943 polypeptide that is bound to said bioactive molecule and allowing said PRO943 and PRO183, PRO184 or PRO185 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

73. The method according to Claim 72, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

74. The method according to Claim 73, wherein said bioactive molecule causes the death of said cell.

75. A method of linking a bioactive molecule to a cell expressing a PRO331 polypeptide, said method comprising contacting said cell with a PRO1133 polypeptide that is bound to said bioactive molecule and allowing said PRO331 and PRO1133 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

76. The method according to Claim 75, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

77. The method according to Claim 75, wherein said bioactive molecule causes the death of said cell.

78. A method of linking a bioactive molecule to a cell expressing a PRO1133 polypeptide, said method comprising contacting said cell with a PRO331 polypeptide that is bound to said bioactive molecule and allowing said PRO331 and PRO1133 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

79. The method according to Claim 78, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

80. The method according to Claim 78, wherein said bioactive molecule causes the death of said cell.

81. A method of linking a bioactive molecule to a cell expressing a PRO1387 polypeptide, said method comprising contacting said cell with a PRO363 or PRO5723 polypeptide that is bound to said bioactive molecule and allowing said PRO1387 and PRO363 or PRO5723 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

82. The method according to Claim 81, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

83. The method according to Claim 81, wherein said bioactive molecule causes the death of said cell.

84. A method of linking a bioactive molecule to a cell expressing a PRO363 or PRO5723 polypeptide, said method comprising contacting said cell with a PRO1387 polypeptide that is bound to said bioactive molecule and allowing said PRO1387 and PRO363 or PRO5723 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

85. The method according to Claim 84, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

86. The method according to Claim 84, wherein said bioactive molecule causes the death of said cell.

87. A method of linking a bioactive molecule to a cell expressing a PRO1114 polypeptide, said method comprising contacting said cell with a PRO3301 or PRO9940 polypeptide that is bound to said bioactive molecule and allowing said PRO1114 and PRO3301 or PRO9940 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

5 88. The method according to Claim 87, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

89. The method according to Claim 87, wherein said bioactive molecule causes the death of said cell.

10 90. A method of linking a bioactive molecule to a cell expressing a PRO3301 or PRO9940 polypeptide, said method comprising contacting said cell with a PRO1114 polypeptide that is bound to said bioactive molecule and allowing said PRO1114 and PRO3301 or PRO9940 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

15 91. The method according to Claim 90, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

20 92. The method according to Claim 90, wherein said bioactive molecule causes the death of said cell.

25 93. A method of linking a bioactive molecule to a cell expressing a PRO1181 polypeptide, said method comprising contacting said cell with a PRO7170, PRO361 or PRO846 polypeptide that is bound to said bioactive molecule and allowing said PRO1181 and PRO7170, PRO361 or PRO846 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

94. The method according to Claim 93, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

30 95. The method according to Claim 93, wherein said bioactive molecule causes the death of said cell.

35 96. A method of linking a bioactive molecule to a cell expressing a PRO7170, PRO361 or PRO846 polypeptide, said method comprising contacting said cell with a PRO1181 polypeptide that is bound to said bioactive molecule and allowing said PRO1181 and PRO7170, PRO361 or PRO846 polypeptides to bind to one another, thereby linking said bioactive molecules to said cell.

97. The method according to Claim 96, wherein said bioactive molecule is a toxin, a radiolabel or an antibody.

98. The method according to Claim 96, wherein said bioactive molecule causes the death of said cell.

99. A method of modulating at least one biological activity of a cell expressing a PRO943 polypeptide, said method comprising contacting said cell with a PRO183, PRO184 or PRO185 polypeptide or an anti-PRO943 antibody, whereby said PRO183, PRO184 or PRO185 polypeptide or said anti-PRO943 antibody binds to said PRO943 polypeptide, thereby modulating at least one biological activity of said cell.

100. The method according to Claim 99, wherein said cell is killed.

101. A method of modulating at least one biological activity of a cell expressing a PRO183, PRO184 or PRO185 polypeptide, said method comprising contacting said cell with a PRO943 polypeptide or an anti-PRO183, anti-PRO184 or anti-PRO185 antibody, whereby said PRO943 polypeptide or said anti-PRO183, anti-PRO184 or anti-PRO185 antibody binds to said PRO183, PRO184 or PRO185 polypeptide, thereby modulating at least one biological activity of said cell.

102. The method according to Claim 101, wherein said cell is killed.

103. A method of modulating at least one biological activity of a cell expressing a PRO331 polypeptide, said method comprising contacting said cell with a PRO1133 polypeptide or an anti-PRO331 antibody, whereby said PRO1133 polypeptide or said anti-PRO331 antibody binds to said PRO331 polypeptide, thereby modulating at least one biological activity of said cell.

104. The method according to Claim 103, wherein said cell is killed.

105. A method of modulating at least one biological activity of a cell expressing a PRO1133 polypeptide, said method comprising contacting said cell with a PRO331 polypeptide or an anti-PRO1133 antibody, whereby said PRO331 polypeptide or said anti-PRO1133 antibody binds to said PRO1133 polypeptide, thereby modulating at least one biological activity of said cell.

106. The method according to Claim 105, wherein said cell is killed.

107. A method of modulating at least one biological activity of a cell expressing a PRO1387 polypeptide, said method comprising contacting said cell with a PRO363 or PRO5723 polypeptide or an anti-PRO1387 antibody, whereby said PRO363 or PRO5723 polypeptide or said anti-PRO1387 antibody binds to said PRO1387 polypeptide, thereby modulating at least one biological activity of said cell.

108. The method according to Claim 107, wherein said cell is killed.

109. A method of modulating at least one biological activity of a cell expressing a PRO363 or PRO5723 polypeptide, said method comprising contacting said cell with a PRO1387 polypeptide or an anti-PRO363 or anti-PRO5723 antibody, whereby said PRO1387 polypeptide or said anti-PRO363 or anti-PRO5723 antibody binds to said PRO363 or PRO5723 polypeptide, thereby modulating at least one biological activity of said cell.

110. The method according to Claim 109, wherein said cell is killed.

111. A method of modulating at least one biological activity of a cell expressing a PRO1114 polypeptide, said method comprising contacting said cell with a PRO3301 or PRO9940 polypeptide or an anti-PRO1114 antibody, whereby said PRO3301 or PRO9940 polypeptide or said anti-PRO1114 antibody binds to said PRO1114 polypeptide, thereby modulating at least one biological activity of said cell.

112. The method according to Claim 111, wherein said cell is killed.

113. A method of modulating at least one biological activity of a cell expressing a PRO3301 or PRO9940 polypeptide, said method comprising contacting said cell with a PRO1114 polypeptide or an anti-PRO3301 or anti-PRO9940 antibody, whereby said PRO1114 polypeptide or said anti-PRO3301 or anti-PRO9940 antibody binds to said PRO3301 or PRO9940 polypeptide, thereby modulating at least one biological activity of said cell.

114. The method according to Claim 113, wherein said cell is killed.

115. A method of modulating at least one biological activity of a cell expressing a PRO1181 polypeptide, said method comprising contacting said cell with a PRO7170, PRO361 or PRO846 polypeptide or an anti-PRO1181 antibody, whereby said PRO7170, PRO361 or PRO846 polypeptide or said anti-PRO1181 antibody binds to said PRO1181 polypeptide, thereby modulating at least one biological activity of said cell.

116. The method according to Claim 115, wherein said cell is killed.

117. A method of modulating at least one biological activity of a cell expressing a PRO7170, PRO361 or PRO846 polypeptide, said method comprising contacting said cell with a PRO1181 polypeptide or an anti-PRO7170, anti-PRO361 or anti-PRO846 antibody, whereby said PRO1181 polypeptide or said anti-PRO7170, anti-PRO361 or anti-PRO846 antibody binds to said PRO7170, PRO361 or PRO846 polypeptide, thereby modulating at least one biological activity of said cell.

5

118. The method according to Claim 117, wherein said cell is killed.

1/330

FIGURE 1

CGGACGCGTGGGTGCGAGGCGAAGGTGACCGGGGACCGAGCATTTTCAGATCTGCTCGGTAGA
CCTGGTGCACCACCACC**ATG**TTGGCTGCAAGGCTGGTGTGTCTCCGGACACTACCTTCTAGG
GTTTTCCACCCAGCTTTCACCAAGGCCTCCCCTGTTGTGAAGAATTCATCACGAAGAATCA
ATGGCTGTTAACACCTAGCAGGGAATATGCCACCAAACAAGAATTGGGATCCGGCGTGGA
GAACTGGCCAAGAAGCTCAAAGAGGCAGCATTTGGAACCATCGATGGAAAAAATATTTAAAT
GATCAGATGGGAAGATGGTTTGTGTGCTGGAGGGGCTGCTGTTGGTCTTGGAGCATTTGTGCTA
CTATGGCTTGGGACTGTCTAATGAGATTGGAGCTATTGAAAAGGCTGTAATTTGGCCTCAGT
ATGTCAAGGATAGAATTCATTCCACCTATATGTACTTAGCAGGGAGTATTGGTTTAACAGCT
TTGTCTGCCATAGCAATCAGCAGAACGCCTGTTCTCATGAACCTCATGATGAGAGGCTCTTG
GGTGACAATTGGTGTGACCTTTGCAGCCATGGTTGGAGCTGGAATGCTGGTACGATCAATAC
CATATGACCAGAGCCCAGGCCCAAAGCATCTTGCTTGGTTGCTACATTCTGGTGTGATGGGT
GCAGTGGTGGCTCCTCTGACAATATTAGGGGGTCTCTCTCATCAGAGCTGCATGGTACAC
AGCTGGCATTTGTGGGAGGCCTCTCCACTGTGGCCATGTGTGCGCCAGTGAAAAGTTTCTGA
ACATGGGTGCACCCCTGGGAGTGGGCCTGGGTCTCGTCTTTGTGTCTCATTTGGGATCTATG
TTTCTTCCACCTACCACCGTGGCTGGTGCCACTCTTTACTCAGTGGCAATGTACGGTGGATT
AGTTCTTTTCAGCATGTTCTTCTGTATGATACCCAGAAAGTAATCAAGCGTGCAGAAGTAT
CACCAATGTATGGAGTTCAAAAATATGATCCCATTAACTCGATGCTGAGTATCTACATGGAT
ACATTAAATATATTTATGCGAGTTGCAACTATGCTGGCAACTGGAGGCAACAGAAAGAA**TG**
AAGTGACTCAGCTTCTGGCTTCTCTGCTACATCAAATATCTTGTTTAATGGGGCAGATATGC
ATTAAATAGTTTGTACAAGCAGCTTTCGTTGAAGTTTAGAAGATAAGAAACATGTCATCATA
TTTAAATGTTCCGGTAATGTGATGCCTCAGGTCTGCCTTTTTTTCTGGAGAATAAATGCAGT
AATCCTCTCCCAAATAAGCACACACATTTTCAATTCTCATGTTTGAGTGATTTTAAATGTT
TTGGTGAATGTGAAAATAAGTTTGTGTGTCATGAGAATGTAAGTCTTTTTTCTACTTTAAAA
TTTAGTAGGTTCACTGAGTAACTAAAATTTAGCAAACCTGTGTTTGCATATTTTTTTGGAGT
GCAGAATATTGTAATTAATGTCATAAGTGATTTGGAGCTTGGTAAAGGGACCAGAGAGAAG
GAGTCACCTGCAGTCTTTTGTTTTTTTAAATACTTAGAACTTAGCACTTGTGTTATTGATTA
GTGAGGAGCCAGTAAGAAACATCTGGGTATTTGGAAACAAGTGGTCATTGTTACATTCATTT
GCTGAACTTAACAAAAGTTCATCCTGAAACAGGCACAGGTGATGCATTCTCCTGCTGTTG
CTTCTCAGTGCTCTCTTTCCAATATAGATGTGGTCATGTTTGACTTGTACAGAATGTTAATC
ATACAGAGAATCCTTGATGGAATTATATATGTGTGTTTTACTTTTGAATGTTACAAAAGGAA
ATAACTTTAAAGTATTCTCAAGAGAAAATATTCAAAGCATGAAATATGTTGCTTTTTCCAG
AATACAAACAGTATACTCATG

2/330

FIGURE 2

MLAARLVCLRTLPSRVFHPAFTKASPVVKNSITKNQWLLTPSREYATKTRIGIRRGRTGQEL
KEAALEPSMEKIFKIDQMGRWFWAGGAAGVGLGALCYYGGLSNEIGAIEKAVIWPQYVKDRI
HSTYMYLAGSIGLTALSAIAISRTPVLMNFMMRGSWVTIGVTFAAMVGAGMLVRSIPYDQSP
GPKHLAWLLHSGVMGAVVAPLTILGGPLLIRAAWYTAGIVGGLSTVAMCAPSEKFLNMGAPL
GVGLGLVFVSSLGSMFLPPTTVAGATLYSVAMYGGLVLFMFLLYDTQKVIKRAEVSPMYGV
QKYDPINSMLSIYMDTLNIFMRVATMLATGGRKK

3/330

FIGURE 3

GAAGGCTGCCTCGCTGGTCCGAATTCGGTGGCGCCACGTCCGCCCCGTCTCCGCCTTCTGCAT
CGCGGCTTCGGCGGCTTCCACCTAGACACCTAACAGTCGCGGAGCCGGCCGCGTCTGAGGG
GGTCGGCACGGGGAGTCGGGCGGTCTTGTGCATCTTGGCTACCTGTGGGTCTGAAG**ATG**TCGG
ACATCGGAGACTGGTTTCAGGAGCATCCCGGCGATCACGCGCTATTGGTTCCGCCGCCACCGTC
GCCGTGCCCTTGGTCGGCAAACCTCGGCCTCATCAGCCCGGCCTACCTCTTCCCTCTGGCCCGA
AGCCTTCCTTTATCGCTTTTCAGATTTGGAGGGCCAATCACTGCCACCTTTTATTTCCCTGTGG
GTCCAGGAACCTGGATTTCTTTATTTGGTCAATTTATATTTCTTATATCAGTATTCTACGCGA
CTTGAAACAGGAGCTTTTGTATGGGAGGCGAGAGACTATTTATTCATGCTCCTCTTTAACTG
GATTTGCATCGTGATTACTGGCTTAGCAATGGATATGCAGTTGCTGATGATTCTCTTGATCA
TGTCAGTACTTTATGTCTGGGCCCAGCTGAACAGAGACATGATTGTATCATTTTTGGTTTGA
ACACGATTTAAGGCCTGCTATTTACCCTGGGTTATCCTTGGATTCAACTATATCATCGGAGG
CTCGGTAATCAATGAGCTTATTGGAAATCTGGTTGGACATCTTTATTTTTTTCCTAATGTTCA
GATACCCAATGGACTTGGGAGGAAGAAATTTCTATCCACACCTCAGTTTTTGTACCGCTGG
CTGCCAGTAGGAGAGGAGGAGTATCAGGATTTGGTGTGCCCCCTGCTAGCATGAGGCGAGC
TGCTGATCAGAATGGCGGAGGCGGGAGACACAACCTGGGGCCAGGGCTTTCGACTTGGAGACC
AG**TGA**AGGGGCGGCCTCGGGCAGCCGCTCCTCTCAAGCCACATTTCTCCAGTGCTGGGTG
CACTTAACAACCTGCGTTCTGGCTAACACTGTTGGACCTGACCCACACTGAATGTAGTCTTTC
AGTACGAGACAAAGTTTCTTAAATCCCGAAGAAAAATATAAGTGTTCCACAAGTTTCACGAT
TCTCATTCAAGTCCTTACTGCTGTGAAGAACAATAACCAACTGTGCAAATTGCAAAACTGAC
TACATTTTTTGGTGTCTTCTCTTCTCCCCTTTCCGTCTGAATAATGGGTTTTAGCGGGTCTT
AATCTGCTGGCATTGAGCTGGGGCTGGGTACCAAACCCCTTCCCAAAGGACCTTATCTCTT
TCTTGCACACATGCCTCTCTCCCACCTTTTCCCAAACCCCCACATTTGCAACTAGAAAAAGTTG
CCCAATAAATTGCTCTGCCCTTGACAGGTTCTGTTATTTATTGACTTTTGCCAAGGCTGGT
ACAACAATCATATGTAGCCCTACTGGCTTTTGGTGGCAGAACTGTTACCAATGAGGGGAG
AAGACAGCCACGGATGAAGCGTTTCTCAGCTTTTGGAAATGCTTCGACTGACATCCGTTGTT
AACCGTTTGCCACTCTTCAGATATTTTTTATAAAAAAAGTACCCTGAGTTCATGAGGGCCA
CAGATTGGTTATTAATGAGATACGAGGGTTGGTGCTGGGTGTTTGTTCCTGAGCTAAGTGA
TCAAGACTGTAGTGAGTTGCAGCTAACATGGGTTAGGTTTAAACCATGGGGGATGCACCCC
TTTGCGTTTTCATATGTAGCCCTACTGGCTTTTGTGTAGCTGGAGTAGTTGGGTGCTTGTGT
TAGGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTCTTTGAGAGGTCCTGGGCATTG
ATTCCCATTTCATCTCATTCTGGATATGTGTTCAATTGAGTAAAGGAGGAGAGACCCTCATA
CGCTATTTAAATGTCACCTTTTTTGCCTATCCCCCGTTTTTTTGGTTCATGTTTCAATTAATTGT
GAGGAAGGCGCAGCTCCTCTCTGCAGTAGATCATTTTTTTAAAGCTAATGTAAGCACATCTA
AGGGAATAACATGATTTAAGGTTGAAATGGCTTTTAGAATCATTGTTGGGTTTGAGGGTGTGTTA
TTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTT
TCGTAGGTGGGCTTTTCCCTATCAGAGCTTGGCTCATAACCAAATAAAGTTTTTTTGAAGGCCA
TGGCTTTTTCACACAGTTATTTTATTTTATGACGTTATCTGAAAGCAGACTGTTAGGAGCAGT
ATTGAGTGGCTGTCACACTTTGAGGCAACTAAAAAGGCTTCAAACGTTTTGATCAGTTTCTT
TTCAGGAAACATTTGTGCTCTAACAGTATGACTATTCTTTCCCCCACTCTTAAACAGTGTGAT
GTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACCTTCTCATTTTGAATAGAGTTTGTGTG
TACTTCTCCATATTTAATTTATATGATAAAATAGGTGGGGAGAGTCTGAACCTTAACTGTCA
TGTTTTGTTGTTTCATCTGTGGCCACAATAAAGTTTACTTGTAAAATTTTAGAGGGCCATTACT
CCAATTATGTTGCACGTACACTCATTGTACAGGCGTGGAGACTCATTGTATGTATAAGAATA
TTTCTGACAGTGAGTGACCCGGAGTCTCTGGTGTACCCCTTTACCAGTCAGCTGCCTGCGAG
CAGTCATTTTTTTCCTAAAGGTTTACAAGTATTTAGAACCTTTTCAGTTCAGGGCAAAATGTTT
ATGAAGTTATTCCTCTTAAACATGGTTAGGAAGCTGATGACGTTATTGATTTTGTCTGGATT
ATGTTTCTGGAATAATTTTACCAAAACAAGCTATTTGAGTTTTGACTTGACAAGGCAAAACA
TGACAGTGGATTCTCTTTACAAATGGAATAAAAAAATCCTTATTTTGTATAAAGGACTTCCC
TTTTTGTAACATAATCCTTTTTTATTGGTAAAAATGTAAATTAATAATGTGCAACTTG

4/330

FIGURE 4

MSDIGDWFRSIPAITRYWFAATVAVPLVGKLGGLISPAYLFLWPEAFLYRFQIWRPITATFYF
PVGPGTGFLYLVNLYFLYQYSTRLETGAFDGRPADYLFMLLFNWICIVITGLAMDMQLLMIP
LIMSVLYVWAQLNRDMIVSFWFGTRFKACYLPWVILGFNYIIIGGSVINELIGNLVGHLYFFL
MFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPASMRRAADQNGGGGRHNWGQGFRL
GDQ

Transmembrane domain:

amino acids 98-116, 152-172

N-myristoylation site.

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

Glycosaminoglycan attachment site.

amino acids 218-222

5/330

FIGURE 5

GGGGCCGCGGTCTAGGGCGGCTACGTGTGTTGCCATAGCGACCATTTTGCATTAACTGGTTG
GTAGCTTCTATCCTGGGGGCTGAGCGACTGCGGGCCAGCTCTTCCCCTACTCCCTCTCGGCT
CCTTGTGGCCCAAAGGCCTAACCGGGGTCCGGCGGTCTGGCCTAGGGATCTTCCCCGTTGCC
CCTTTGGGGCGGG**ATG**GCTGCGGAAGAAGAAGACGAGGTGGAGTGGGTAGTGGAGAGCATCG
CGGGGTTCCCTGCGAGGCCAGACTGGTCCATCCCCATCTTGGACTTTGTGGAACAGAAATGT
GAAGTTAACTGCAAAGGAGGGCATGTGATAACTCCAGGAAGCCCAGAGCCGGTGATTTTGGT
GGCCTGTGTTCCCCTTGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGATTC
ATCAGGAATACAAAGAACTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAATT
AATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAGGC
CATTTTGCACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCCAGA
AAAACATTGAAATGCAGCTGCAAGCCATTGCAATAATTCAAGAGAGAAATGGTGTATTACCT
GACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAATCCT
GAGGGAAGTTCTTAGAAAATCAAAAGAGGAATATGACCAGGAAGAAGAAAGGAAGAGGAAAA
AACAGTTATCAGAGGCTAAAACAGAAGAGCCACAGTGCATTCCAGTGAAGCTGCAATAATG
AATAATTCCCAAGGGGATGGTGAACATTTTGCACACCCACCCTCAGAAGTTAAAATGCATTT
TGCTAATCAGTCAATAGAACCTTTGGGAAGAAAAGTGGAAAGGTCTGAAACTTCCTCCCTCC
CACAAAAGGCCTGAAGATTCCCTGGCTTAGAGCATGCGAGCATTGAAGGACCAATAGCAAAC
TTATCAGTACTTGGAACAGAAGAACTTCGGCAACGAGAACACTATCTCAAGCAGAAGAGAGA
TAAGTTGATGTCCATGAGAAAGGATATGAGGACTAAACAGATACAAAATATGGAGCAGAAAG
GAAAACCCACTGGGGAGGTAGAGGAAATGACAGAGAAACCAGAAATGACAGCAGAGGAGAAG
CAAACATTACTAAAGAGGAGATTGCTTGCAAGAGAACTCAAAGAAGAAGTTATTAATAAG**TA**
ATAATTAAGAACAATTTAACAAAATGGAAGTTCAAATTGTCTTAAAAATAAATTATTTAGTC
CTTACACTG

6/330

FIGURE 6

MAAEEDEVEWVVESIAGFLRGPDWSIPILDFVEQKCEVNCKGGHVITPGSPEPVILVACVP
LVFDDEEESKLTYTEIHQEYKELVEKLLEGYLKEIGINEDQFQEACTSPLAKTHTSQAILQP
VLAAEDFTIFKAMMVQKNIEMQLQAIRIIQERNGVLPDCLTDGSDVSDLEHEEMKILREVL
RKSKEEYDQEEERKRKKQLSEAKTEEPTVHSSEAAIMNNSQGDGEHFAHPPSEVKMHFANQS
IEPLGRKVERSETSSLPQKGLKIPGLEHASIEGPIANLSVLGTEELRQREHYLKQKRDKLMS
MRKDMRTKQIQNMEQKGKPTGEVEEMTEKPEMTAEKQTLLKRLLAEKLKEEVINK

N-glycosylation sites.

amino acids 224-228, 246-250, 285-289

N-myristoylation site.

amino acids 273-279

Amidation site.

amino acids 252-256

Cytosolic fatty-acid binding proteins.

amino acids 78-108

7/330

FIGURE 7

GGGCACAGCACATGTGAAGTTTTTGATGATGAAGAAGAAAGCAAATTGACCTATACAGAGAT
TCATCAGGAATACAAAGAAGCTAGTTGAAAAGCTGTTAGAAGGTTACCTCAAAGAAATTGGAA
TTAATGAAGATCAATTTCAAGAAGCATGCACTTCTCCTCTTGCAAAGACCCATACATCACAG
GCCATTTTTGCAACCTGTGTTGGCAGCAGAAGATTTTACTATCTTTAAAGCAATGATGGTCC
AGAAAAACATTGAAATGCAGCTGCAAGCCATTCGAATAATTCAAGAGAGAAATGGTGTATTA
CCTGACTGCTTAACCGATGGCTCTGATGTGGTCAGTGACCTTGAACACGAAGAGATGAAAAT
CCTGAGGGAAGTTCTTAGAAAATCAAAGAGGAATATGACCAGGAA

8/330

FIGURE 8

GCGTGGTTTTTGTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTTCGCCTATACCTACTG
TAGCTTCTCCACGTATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTG
TCTCAGCTCTAGGATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTA AAAAAC
AGTGGAAATGAAAAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATAC
ATTCTGCTAGGTGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATT
CTGCCAATGAAGAAAACAAGTATGATTATCTTCCAACCTACTGTGAATGTGTGCTCAGAACTG
GTGAAGCTAGTTTTCTGTGTGCTTGTGTCAATCTGTGTTATAAAGAAAGATCATCAAAGTAG
AAATTTGAAATATGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTCTGCCT
TTCTTTATTTTCTGGATAACTTGATTGTCTTCTATGTCTGTCTATCTTCAACCAGCCATG
GCTGTTATCTTCTCAAATTTTAGCATTATAACAACAGCTCTTCTATTCAAGGATAGTGCTGAA
GAGGCGTCTAAACTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTTGTCTATTGTGGCCT
TGACTGCCGGGACTAAACTTTACAGCACAACTTGGCAGGACGTGGATTTTCATCACGATGCC
TTTTTTCAGCCCTTCCAATTCCTGCCTTCTTTTCAGAAGTGAGTGTCCAGAAAAGACAATTG
TACAGCAAAGGAATGGACTTTTCTGAAGCTAAATGGAACACCACAGCCAGAGTTTTTCAGTC
ACATCCGTCTTGGCATGGGCCATGTTCTTATTATAGTCCAGTGTTTTATTTCTTCAATGGCT
AATATCTATAATGAAAAGATACTGAAGGAGGGGAACCAGCTCACTGAAAGCATCTTCATACA
GAACAGCAAACCTCTATTTCTTTGGCATTCTGTTTAATGGGCTGACTCTGGGCCTTCAGAGGA
GTAACCGTGATCAGATTAAGAACTGTGGATTTTTTTTATGGCCACAGTGCATTTTCAGTAGCC
CTTATTTTTTGTAACTGCATTCCAGGGCCTTTCAGTGGCTTTCATTCTGAAGTTCCTGGATAA
CATGTTCCATGTCTTGATGGCCCAGGTACCAGTGTCAATTATCACAACAGTGTCTGTCTCTGG
TCTTTGACTTCAGGCCCTCCCTGGAATTTTTCTTGAAGCCCCATCAGTCCTTCTCTCTATA
TTTATTTTATAATGCCAGCAAGCCTCAAGTTCGGGAATACGCACCTAGGCAAGAAAGGATCCG
AGATCTAAGTGGCAATCTTTGGGAGCGTTCCAGTGGGGATGGAGAAGAACTAGAAAGACTTA
CCAAACCCAAGAGTGATGAGTCAGATGAAGATACTTCTTAACTGGTACCCACATAGTTTTGCA
GCTCTCTTGAACCTTATTTTTCACATTTTTCAGTGTTTGTAATATTTATCTTTTCACTTTGATA
AACCAGAAATGTTTCTAAATCCTAATATTCTTTGCATATATCTAGCTACTCCCTAAATGGTT
CCATCCAAGGCTTAGAGTACCCAAAGGCTAAGAAATTCTAAAGAACTGATACAGGAGTAACA
ATATGAAGAATTTCATTAATATCTCAGTACTTGATAAATCAGAAAGTTATATGTGCAGATTAT
TTTCTTGGCCTTCAAGCTTCCAAAAAACTTGTAATAATCATGTTAGCTATAGCTTGTATAT
ACACATAGAGATCAATTTGCCAAATATTCACAATCATGTAGTTCTAGTTTACATGCCAAAGT
CTTCCCTTTTTTAACATTATAAAAGCTAGGTTGTCTCTTGAATTTTGAGGCCCTAGAGATAGT
CATTTTGCAAGTAAAGAGCAACGGGACCCTTTCTAAAAACGTTGGTTGAAGGACCTAAATAC
CTGGCCATACCATAGATTTGGGATGATGTAGTCTGTGCTAAATATTTTGTCTGAAGAAGCAGT
TTCTCAGACACAACATCTCAGAATTTTAATTTTTTAGAAATTCATGGGAAATTGGATTTTTGT
AATAATCTTTTGATGTTTTTAAACATTGGTTCCCTAGTCACCATAGTTACCCTTGTATTTTA
AGTCATTTTAAACAAGCCACGGTGGGGCTTTTTTCTCCTCAGTTTGAGGAGAAAAATCTTGAT
GTCATTACTCCTGAATTATTACATTTTGGAGAATAAGAGGGCATTTTATTTTATTAGTTACT
AATTCAAGCTGTGACTATTGTATATCTTTCCAAGAGTTGAAATGCTGGCTTCAGAATCATAAC
CAGATTGTCAGTGAAGCTGATGCCTAGGAACTTTTAAAGGGATCCTTTCAAAGGATCACTT
AGCAAACACATGTTGACTTTTAACTGATGTATGAATATTAATACTCTAAAAATAGAAAGACC
AGTAATATATAAGTCACCTTACAGTGCTACTTCACACTTAAAGTGCATGGTATTTTTTCATG
GTATTTTGCATGCAGCCAGTTAACTCTCGTAGATAGAGAAGTCAGGTGATAGATGATATTAA
AAATTAGCAAACAAAAGTGACTTGCTCAGGGTCATGCAGCTGGGTGATGATAGAAGAGTGGG
CTTTAACTGGCAGGCCTGTATGTTTACAGACTACCATACTGTAAATATGAGCTTTATGGTGT
CATTCTCAGAACTTATACATTTCTGCTCTCCTTTCTCCTAAGTTTCATGCAGATGAATATA
AGGTAATATACTATTATATAATTCATTTGTGATATCCACAATAATATGACTGGCAAGAATTG
GTGGAAATTTGTAATTAAATAATTATTAAACCT

9/330

FIGURE 9

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVK
LVFCVLVSFCVIKKDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLQPAMAV
IFS NFSIIT TALLFRIVLKRRLNWIQWASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFF
SPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFSHIRLGMGHVLIIVQCFISSMANI
YNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFYGHS AFSVALI
FVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIIITTVSVLVFDFRPSLEFFLEAPSVLLSIFI
YNASKPQVPEYAPRQERIRDL SGNLWERSSSGDGEELERLTKPKSDESD EDTF

Transmembrane domains:

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293,
298-318, 328-368

N-glycosylation sites.

amino acids 128-132, 204-208, 218-222, 374-378

Glycosaminoglycan attachment site.

amino acids 402-406

N-myristoylation sites.

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

10/330

FIGURE 10

CGTGCCTGCGCAATGGGTGTCGGGTCCGCTTTTTTCCCAATCCGGACGTAATCGTGGTTTTTG
TTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTTCGCCTATACCTACTGTAGCTTCTCCAC
GTATGGACCCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTGTCTCAGCTCTAG
GATGTGCGTTCTTCCACTAGAAGCTCTTCTGAGGGAGGTAATTAAAAACAGTGGAATGGAA
AAACAGTGCTGTAGTCATCCTGTAATATGCTCCTTGTCAACAATGTATACATTCCCTGCTAGG
TGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTAGTGAAGTATTCTGCCAATGAAG
AAAACAAGTATGATTATCTTCCAACACTACTGTGAATGTGTGCTCAGAACTGGTGAAGCTAGTT
TTCTGTGTGCTTGTGTCATTCTGTGTTATAAAGAAAGATCATCAAAGTAGAAATTTGAAATA
TGCTTCCTGGAAGGAATTCTCTGATTTTCATGAAGTGGTCCATTCCCTGCCTTTCTTTATTTCC
TGGATAACTTGATTGTCTTCTATGTCCTGTCCTATCTTCAACCAGCCATGGCTGTTATCTTC
TCAAATTTTAGCATTATAACAACAGCTCTTCTATTTCAGGATAGTGCTGAAGAGGCGTCTAAA
CTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATTGTGGCCTTGACTGCCGGGA
CTAAACTTTA

11/330

FIGURE 11

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGGCCGGCTTGGCTAGCGCGCGGGCGGCC
GTGGCTAAGGCTGCTACGAAGCGAGCTTGGGAGGAGCAGCGGCCTGCGGGGCAGAGGAGCAT
CCCGTCTACCAGGTCCCAAGCGGCGTGGCCCGCGGGTCATGGCCAAAGGAGAAGGCGCCGAG
AGCGGCTCCGCGGGCGGGGCTGCTACCCACCAGCATCCTCCAAAGCACTGAACGCCCGGCCCA
GGTGAAGAAAGAACCAGAAAAAGAAACAACAGTTGTCTGTTTGCAACAAGCTTTGCTATG
CACTTGGGGGAGCCCCCTACCAGGTGACGGGCTGTGCCCTGGGTTTCTTCCCTTCAGATCTAC
CTATTGG**ATG**TGGCTCAGGTGGGCCCTTTCTCTGCCTCCATCATCCTGTTTGTGGGCCGAGC
CTGGGATGCCATCACAGACCCCCTGGTGGGCCTCTGCATCAGCAAATCCCCCTGGACCTGCC
TGGGTGCCTTATGCCCTGGATCATCTTCTCCACGCCCCTGGCCGTCATTGCCTACTTCCTC
ATCTGGTTCGTGCCCCGACTTCCCACACGGCCAGACCTATTGGTACCTGCTTTTCTATTGCCT
CTTTGAAACAATGGTCACGTGTTTCCATGTTCCCTACTCGGCTCTCACCATGTTTCATCAGCA
ACCGAGCAGACTGAGCGGGATTCTGCCACCGCCTATCGGATGACTGTGGAAGTGCTGGGCAC
AGTGCTGGGCACGGCGATCCAGGGACAAATCGTGGGCCAAGCAGACACGCCTTGTTTCCAGG
ACTTCAATAGCTCTACAGTAGCTTCACAAAGTGCCAACCATAACATGGCACCACCTTCACAC
AGGGAAACGCAAAGGCATACCTGCTGGCAGCGGGGGTCAATTGTCTGTATCTATATAATCTG
TGCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTG
AGCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATAACATCAAACCT
ATTACTGGCTTCCTCTTCACCTCCTTGGCTTTTCATGCTGGTGGAGGGGAACCTTTGTCTTGT
TTGCACCTACACCTTGGGCTTCCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCT
CGGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCT
GTATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAA
CCTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTAC
TACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCAT
GGAACCGAGCCCATCTTCTTCTCCTTCTATGTCTTCTTCACCAAGTTTGCCTCTGGAGTGTC
ACTGGGCATTTCTACCCTCAGTCTGGACTTTGCAGGGTACCAGACCCGTGGCTGCTCGCAGC
CGGAACGTGTCAAGTTTACACTGAACATGCTCGTGACCATGGCTCCCATAGTTCTCATCCTG
CTGGGCCTGCTGCTCTTCAAATGTACCCCATTTGATGAGGAGAGGCGGCGGCAGAATAAGAA
GGCCCTGCAGGCACTGAGGGACGAGGCCAGCAGCTCTGGCTGCTCAGAAACAGACTCCACAG
AGCTGGCTAGCATCCTC**TAG**GGCCCGCCACGTTGCCCGAAGCCACCATGCAGAAGGCCACAG
AAGGGATCAGGACCTGTCTGCCGGCTTGCTGAGCAGCTGGACTGCAGGTGCTAGGAAGGGAA
CTGAAGACTCAAGGAGGTGGCCCAGGACACTTGCTGTGCTCACTGTGGGGCCGGCTGCTCTG
TGGCCTCCTGCCTCCCCTCTGCCTGCCTGTGGGGCCAAGCCCTGGGGCTGCCACTGTGAATA
TGCCAAGGACTGATCGGGCCTAGCCCGGAACACTAATGTAGAAACCTTTTTTTTACAGAGCC
TAATTAATAACTTAATGACTGTGTACATAGCAATGTGTGTGTATGTATATGTCTGTGAGCTA
TTAATGTTATTAATTTTCATAAAAGCTGGAAAGC

12/330

FIGURE 12

MWLRWALSLPPSSCLWAEPGMPSQTPWWASASANPPGPAWVALCPGSSSPRPWPSLPTSSSG
SCPTSHTARPIGTCFSLKQWSRVSMFPTRLSPCSSATEQTERDSATAYRMTVEVLGTVL
GTAIQQQIVGQADTPCFQDFNSSTVASQSANHHTGTTSHRETQKAYLLAAGVIVCIYIICAV
ILILGVREQREPYEAQQSEPIAYFRGLRLVM SHGPYIKLITGFLFTSLAFMLVEGNFVLFCT
YTLGFRNEFQNL LLAIMLSATLTIPWQWFLTRFGKKTAVYVGISSAVPFLILVALMESNLI
ITYAVAVAAGISVAAAFLLPWSMLPDVIDDFHLKQPHFHGTEPIFFSFYVFFTKFASGVSLG
ISTLSLDFAGYQTRGCSQPERVKFTLNMLVTMAPIVLILLGLLLFKMYPIDEERRRQNKAL
QALRDEASSSGCSETDSTELASIL

13/330

FIGURE 13

GGGAAACGCAAAAGGCATACCTGCTGGCAGCGGGGGTCATTGTCTGTATCTATATAATCTGT
GCTGTCATCCTGATCCTGGGCGTGCGGGAGCAGAGAGAACCCTATGAAGCCCAGCAGTCTGA
GCCAATCGCCTACTTCCGGGGCCTACGGCTGGTCATGAGCCACGGCCCATACATCAAACCTTA
TTACTGGCTTCCTCTTCACCTCCTTGGCTTTCATGCTGGTGGAGGGGAACTTTGTCTTGTTT
TGCACCTACACCTTGGGCTTCGCAATGAATTCCAGAATCTACTCCTGGCCATCATGCTCTC
GGCCACTTTAACCATTCCCATCTGGCAGTGGTTCTTGACCCGGTTTGGCAAGAAGACAGCTG
TATATGTTGGGATCTCATCAGCAGTGCCATTTCTCATCTTGGTGGCCCTCATGGAGAGTAAC
CTCATCATTACATATGCGGTAGCTGTGGCAGCTGGCATCAGTGTGGCAGCTGCCTTCTTACT
ACCCTGGTCCATGCTGCCTGATGTCATTGACGACTTCCATCTGAAGCAGCCCCACTTCCATG
GAACCGAGCCCAT

14/330

FIGURE 14

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGT
ATGAGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAA**ATGT**
GGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCT
TTCATATTTTCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATAT
CAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGG
CAGTTTTATGCATTGCTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAA
GAGAACGTTATCATCAAATTAAACAAGGCTGGCCTTGTAAGTCTGGAATACTGAGTTGTTTAGG
ACTTTCTATTGTGGCAAACCTTCAGAAAAACAACCCTTTTTGCTGCACATGTAAGTGGAGCTG
TGCTTACCTTTGGTATGGGCTCATTATATATGTTTGTTCAGACCATCCTTTTCTACCAAATG
CAGCCCCAAAATCCATGGCAAACAAGTCTTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGG
AGTAAGTGCACCTTAGCATGCTGACTTGCTCATCAGTTTTGCACAGTGGCAATTTTGGGACTG
ATTTAGAACAGAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACT
ACTGCAGCAGAAATGGTCTATGTCATTTTCCTTCCTTGGTTTTTTTCTGACTTACATTCGTGA
TTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACATGGATTAACCCTCTATGACACTG
CACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCAGAGATATTT**TGAT**GAAAGGAT
AAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGGTTACAGAAAGTTGCTTA
TTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCTGATA
ATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTA
TTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACTATG

15/330

FIGURE 15

MWWFQQGLSFLPSALVIWTSAAFIIFSYYITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNI
AAVLCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSG
AVLTFGMGSLYMFVQTILSYQMQPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFG
TDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFLLTYIRDFQKISLRVEANLHGLTLYD
TAPCPINNERTRLLSRDI

16/330

FIGURE 16

CGGACGCTTGGGCNGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGTGCCTGATGCCGAGT
TCCGTCTCTCGGGTCTTTTCCTGGTCCCAGGCAAAGCGGAGCGGAGATCCTCAAACGGCCTA
GTGCTTCGCGCTTCCGGAGAAAATCAGCGGTCTAATTAATTCCTCTGGTTTGTTGAAGCAGT
TACCAAGAATCTTCAACCCTTTCCCACAAAAGCTAATTGAGTACACGTTTCCTGTTGAGTACA
CGTTCCTGTTGATTTACAAAAGGTGCAGGTATGAGCAGGTCTGAAGACTAACATTTTGTGAA
GTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGTTTCAGCAAGGCCTCAGTTTCCTTCCT
TCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTTTCATACATTACTGCAGTAACACT
CCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGGTACAGTANC

17/330

FIGURE 17

CCCACGCGTCCGCCCCGCGCTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCCGCGGG
CCGGGGTGCGGAGCCGAC**ATG**CGCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGCTGCAC
CTTCGCCTTGTACTTGCTGTGACGCGACTGCCCCGCGGGCGGAGACTGGGCTCCACCGAGG
AGGCTGGAGGCAGGTGCTGTGGTTCCCCCTCCGACCTGGCAGAGCTGCGGGAGCTCTCTGAG
GTCCTTCGAGAGTACCGGAAGGAGCACCAGGCCTACGTGTTCCCTGCTCTTCTGCGGCGCCTA
CCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCCTGAATGTTTTAGCTGGTGCCT
TGTTTGGGCCATGGCTGGGGCTTCTGCTGTGCTGTGTGTGACCTCGGTGGGTGCCACATGC
TGCTACCTGCTCTCCAGTATTTTTGGCAAACAGTTGGTGGTGTCTACTTTCCTGATAAAGT
GGCCCTGCTGCAGAGAAAGGTGGAGGAGAACAGAAACAGCTTGTTTTTTTTCTTATTGTTTT
TGAGACTTTTCCCCATGACACCAAAGTGGTTCTTGAACCTCTCGGCCCCAATTCTGAACATT
CCCATCGTGCAGTTCTTCTTCTCAGTTCTTATCGGTTTGATCCCATATAATTCATCTGTGT
GCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTTCTCCTGGGACACTG
TCTTTAAGCTGTTGGCCATTGCCATGGTGGCATTAATTCCTGGAACCCTCATTAAAAAATTT
AGTCAGAAACATCTGCAATTGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAGA
CACAT**TGA**TCTGGATTTTCTGTTTGCCACATCCCTGGACTCAGTTGCTTATTTGTGTAATGGA
TGTGGTCCTCTAAAGCCCCTCATTGTTTTTGATTGCCTTCTATAGGTGATGTGGACACTGTG
CATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATCAGGT
TTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTGT
GGCCGGGCGCGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGTGATTC
ACAAGGTCAGGAGTTCAAGACCAGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAT
ACAAAAATTAGCCAGGCGTGGTGGCAGGCACCTGTAATCCCAGCTACTCGGGAGGCTGAGGC
AGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGATCACACCACTGCACT
CCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

18/330

FIGURE 18

MRPLLGLLLVFAGCTFALYLLSTRLPGRRLGSTEEAGGRSLWFPSDLAELRELVREYR
KEHQAYVFLLFCGAYLYKQGFAIPGSSFLNVLAGALFGPWLGLLLCCVLTSVGATCCYLLSS
IFGKQLVVSYPDPKVALLQKVEENRNSLFFFLFLRLFPMTPNWFLNLSAPILNIPIVQFF
FSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVFKLLAIAMVALIPGTLIKKFSQKHLQ
LNETSTANHIHSRKDT

Important features:**Signal peptide:**

amino acids 1-17

Transmembrane domains:

amino acids 101-123, 189-211

N-glycosylation sites.

amino acids 172-176, 250-254

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 240-244, 261-265

N-myristoylation site.

amino acids 13-19, 104-110, 115-121, 204-210

Amidation site.

amino acids 27-31

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 4-15

Protein splicing proteins.

amino acids 25-31

Sugar transport proteins.

amino acids 162-172

19/330

FIGURE 19

CCGAGGCGGGAGGAGCCCCGAGGGGGCGCGAGCCCCGCATGAATCATTGTAGTCAATCATTTT
CCAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTCCGAACTGCCAGCTCAGAA
TAGGAAAATAACTTGGGATTTTATATTGGAAGAC**ATG**GATCTTGCTGCCAACGAGATCAGCA
TTTATGACAAACTTTCAGAGACTGTTGATTTGGTGAGACAGACCGGCCATCAGTGTGGCATG
TCAGAGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACC
CCCCCGCAGTATCCTCTCCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCT
TGCTCACTGCCTACTTTGTGATTCAACCTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTCT
GGAGCTCACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCCTTGCCCATTGCCAA
GAAGTACATGTCAGAAAATAAGGGAGTTCCTCTGCATGGGGGTGATGAAGACAGACCCTTTC
CAGACTTTGACCCCTGGTGGACAAACGACTGTGAGCAGAATGAGTCAGAGCCCATTCTGCC
AACTGCACTGGCTGTGCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAA
ATTTGAGAGGCTCCATCCACTGGTGATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTC
AGCATTTTTTGTGCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTCGCCAAG
TGGTGGCGCTGCTTTCCTGAGCGGTGGTTCCCATTTCCTTATCCATGGAGGAGACCTCTGAA
CAGATCACAAATGTTACGTGAGCTTTTTCTGTTTTCACTCACCTGCCATTTCCAAAAGATG
CCTCTTTAAACAAGTGCTCCTTTCCTCACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAG
ATGCCTGACCTATTTATCATTGGCAGCGGTGAGGCCATGTTGCAGCTCATCCCTCCCTTCCA
GTGCCGAAGACATTGTCAGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTCTG
ACACCACCCACTGGAAGGTCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGAT
GGAACCGCTTTCTCAGAACTG**TAG**GAAATAGAACTGTGCACAGGAACAGCTTCCAGAGCCGA
AAACCAGGTTGAAAGGGGAAAAATAAAAACAAAAACGATGAAACTGCAAAAA

20/330

FIGURE 20

MDLAANEISIIYDKLSETVDLVRQTGHQCGMSEKAIEKFIRQLLEKNEPQRPPPQYPLLIVVY
KVLATLGLILLTAYFVIQPFSPLAPEPVLGAHTWRS LIHHIRLMSLP IAKKYMSENKGVPL
HGGDEDRPFPDFDPWWTNDCEQNESEPI PANCTGCAQKHLKVMLLEDAPRKFERLHPLVIKT
GKPLLEEEIQHFLCQYPEATEGFSEGFFAKWWRCFPERWFPPYPWRRPLNRSQMLRELFV
FTHLPFPKDASLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQCRRHCSVAMP
IEPGDIGYVDTTHWKVYVIARGVQPLVICDGTAFSEL

21/330

FIGURE 21

CCACGGTGTCCGTTCTTCGCCCCGGCGGCAGCTGTCCCCGAGGCGGGAGGAGCCCCGAGGGGCG
CGAGCCCCGCATGAATCATTGTAGTCAATCATTTTCCAGTTCTCAGCCGTTTCAGTTGTGATC
AAGGGACACGTGGTTTCCGAACTGCCAGCTCAGAATAGGAAAATAACTTGGGATTTTATATT
GGAAGACATGGATCTTGCTGCCAACGAGATCAGCATTTATGACAACTTTCAGAGACTGTTG
ATTTGGTGAGACAGACCGGCCATCAGTGTGGCATGTCAGAGAAGGCAATTGAAAAATTTATC
AGACAGCTGCTGGAAAAGAATGAACCTCAGAGACCCCCCCCCGCAGTATCCTCTCCTTATAGT
TGTGTATAAGGTTCTCGCAACCTTGGGATTAATCTTGCTCACTGCCTACTTTGTGATTCAAC
CTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTGTGGAGCTCAC

22/330

FIGURE 22

CCCACGCGTCCGCCCACGCGTCCGGCTGAACACCTCTTCTTTGGAGTCAGCCACTGATGAGG
CAGGGTCCCCACTTGCAGCTGCAGCAGCTGCAGCAGCTGCAGAGCGCTGCTCCTGGCTGGTG
CCACTGGTGCACGCTGCTAGACCGTGCCTATGAGCCGCTGGGGCTGCAGTGGGGACTGCC
CTCCCTGCCACCCACCAATGGCAGCCCCACCTTCTTTGAAGACTTCCAGGCTTTTTGTGCCA
CACCCGAATGGCGCCACTTCATCGACAAACAGGTACAGCCAACC**ATG**TCCAGTTCGAAATG
GACACGTATGCTAAGAGCCACGACCTTATGTCAGGTTTCTGGAATGCCTGCTATGACATGCT
TATGAGCAGTGGGCAGCGGCCAGTGGGAGCGCGCCAGAGTCGTCGGGCCCTTCCAGGAGC
TGGTGCTGGAACCTGCGCAGAGGGCGGCGCCTGGAGGGGCTACGCTACACGGCAGTGTG
AAGCAGCAGGCAACGCAGCACTCCATGGCCCTGCTGCAGTGGGGGGCGCTGTGGCGCCAGCT
CGCCAGCCCATGTGGGGCCTGGGCGCTGAGGGACACTCCCATCCCCCGCTGGAAACTGTCCA
GCGCCGAGACATATTACGCATGCGTCTGAAGCTGGTGGCCAACCATCACTTCGACCCTCAC
CTGGAAGCCAGCGCTCTCCGAGACAATCTGGGTGAGGTTCCCCTGACACCCACCGAGGAGGC
CTCACTGCCTCTGGCAGTGACCAAAGAGGCCAAAGTGAGCACCCACCCAGTGTGCTGCAGG
AGGACCAGCTCGGCGAGGACGAGCTGGCTGAGCTGGAGACCCCGATGGAGGCAGCAGAACTG
GATGAGCAGCGTGAGAAGCTGGTGCTGTGCGCCGAGTGCCAGCTGGTGACGGTAGTGGCCGT
GGTCCCAGGGCTGCTGGAGGTCACCACACAGAATGTATACTTCTACGATGGCAGCACTGAGC
GCGTGGAACCCGAGGAGGGCATCGGCTATGATTTCCGGCGCCCACTGGCCAGCTGCGTGAG
GTCCACCTGCGGCGTTTCAACCTGCGCGCTTACGACCTTGAAGCTTCTTCTTATCGATCAGGC
CAACTACTTCTCAACTTCCCATGCAAGGTGGGCACGACCCAGTCTCATCTCCTAGCCAGA
CTCCGAGACCCAGCCTGGCCCCATCCCACCCCATACCCAGGTACGGAACCAGGTGTACTCG
TGGCTCCTGCGCCTACGGCCCCCTCTCAAGGCTACCTAAGCAGCCGCTCCCCCAGGAGAT
GCTGCGTGCCTCAGGCCTTACCCAGAAATGGGTACAGCGTGAGATATCCAACTTCGAGTACT
TGATGCAACTCAACACCATTTGCGGGCGGACCTACAATGACCTGTCTCAGTACCCTGTGTTC
CCCTGGGTCTGCGAGGACTACGTGTCCCCAACCCCTGGACCTCAGCAACCAGCCGTCTTCCG
GGACCTGTCTAAGCCCATCGGTGTGGTGAACCCCAAGCATGCCCAGCTCGTGAGGGAGAAGT
ATGAAAGCTTTGAGGACCCAGCAGGGACCATTGACAAGTTCCACTATGGCACCCACTACTCC
AATGCAGCAGGCGTGATGCACTACCTCATCCGCGTGGAGCCCTTCACCTCCCTGCACGTCCA
GCTGCAAAGTGGCCGCTTTGACTGCTCCGACCGGCAGTTCCACTCGGTGGCGGCAGCCCTGGC
AGGCACGCTGGAGAGCCCTGCCGATGTGACAGATGAGCGGGAACCGGAAGGCTCTGGAGGGCATTA
TCAGCAACTTTGGGCAGACTCCCTGTGAGCTGCTGAAGGAGCCACATCCAACTCGGCTCTCA
GCTGAGGAAGCAGCCCATCGCCTTGCACGCCTGGACACTAACTCACCTAGCATCTTCCAGCA
CCTGGACGAACCTCAAGGCATTCTTCGCAGAGGTGACTGTGAGTGCCAGTGGGCTGCTGGGCA
CCCACAGCTGGTTGCCCTATGACCGCAACATAAGCAACTACTTCAGCTTCAGCAAAGACCCC
ACCATGGGCAGCCACAAGACGCAGCAGCTGCTGAGTGGCCCGTGGGTGCCAGGCAGTGGTGT
GAGTGGACAAGCACTGGCAGTGGCCCCGGATGGAAAGCTGCTATTACGCGGTGGCCACTGGG
ATGGCAGCCTGCGGGTGACTGCACTACCCCGTGGCAAGCTGTTGAGCCAGCTCAGCTGCCAC
CTTGATGTAGTAACCTGCCTTGCAGTGGACACCTGTGGCATCTACCTCATCTCAGGCTCCCG
GGACACCACGTGCATGGTGTGGCGGCTCCTGCATCAGGGTGGTCTGTGCTAGTAGGCCTGGCAC
CAAAGCCTGTGCAGGTCTGTATGGGCATGGGGCTGCAGTGAGCTGTGTGGCCATCAGCACT
GAACTTGACATGGCTGTGTCTGGATCTGAGGATGGAAGTGTGATCATAACACTGTACGCCG
CGGACAGTTTGTAGCGGCACTACGGCCTCTGGGTGCCACATTCCCTGGACCTATTTTCCACC
TGGCATTGGGGTCCGAAGGCCAGATTGTGGTACAGAGCTCAGCGTGGGAACGTCTGGGGCC
CAGGTACCTACTCCTTGCACCTGTATTCACTCAATGGGAAGTTGCGGGCTTCACTGCCCT
GGCAGAGCAGCCTACAGCCCTGACGGTGACAGAGGACTTTGTGTTGTGTTGGGACCCGCCAGT
GCGCCCTGCACATCCTCCAACATAACACACTGCTCCCGGCCGCGCCTCCCTTGCCCATGAAG
GTGGCCATCCGCAGCGTGGCCGTGACCAAGGAGCGCAGCCACGTGCTGGTGGGCCCTGGAGGA
TGGCAAGCTCATCGTGGTGGTTCGCGGGGCGAGCCCTCTGAGGTGCGCAGCAGCCAGTTCGCGC
GGAAGCTGTGGCGGTCTCGCGGCGCATCTCCAGGTGTCCTCGGGAGAGACGGAATAACAAC
CCTACTGAGCGCGCT**TGA**ACCTGGCCAGTCCCGGCTGCTCGGGCCCCGGCCCCGGCAGCTG
GCCCCGGGAGGGCCCCGCCCCAGAAGTCGGCGGGGAACACCCCGGGGTGGGCAGCCAGGGGGTGA
GCGGGGCCCCACCTGCCCAGCTCAGGGATTGGCGGGCGATGTTACCCCTCAGGGATTGGCG
GGCGGAAGTCCCGCCCCCTCGCCGGCTGAGGGGCCGCCCTGAGGGCCAGCACTGGCGTCT

23/330

FIGURE 23

MSQFEMDTYAKSHDLMSGFWNACYDMLMSSGQRRQWERAQSRRAFQELVLEPAQRRARLEGL
RYTAVLKQQATQHSMALLHWGALWRQLASPCGAWALRDTPIPRWKLSSAETYSRMRLKLVPN
HHFDPHLEASALRDNLGEVPLTPTEEASLPLAVTKEAKVSTPPELLQEDQLGEDELALETP
MEAAELDEQREKLVLSAECQLVTVVAVVPGLLEVTTQNVYFYDGSTERVETEEGIGYDFRRP
LAQLREVHLRRFNLRRSALELFFIDQANYFLNFPCKVGTTPVSSPSQTPRPQPGPIPPHTQV
RNQVYSWLLRLRPPSQGYLSSRSPQEMLRASGLTQKWVQREISNFEYLMQLNTIAGR TYNDL
SQYPVFPWVLQDYVSPTLDLSNPAVFRDLSPKIGVVNPKHAQLVREKYESFEDPAGTIDKFH
YGTHYSNAAGVMHYLIRVEPFTSLHVQLQSGRFDCSDRQFHSVAAAWQARLESPADV KELIP
EFFYFPDFLENQNGFDLGCLQLTNEKVGDVVLPPWASSPEDFIQQHRQALESEYVSAHLHEW
IDLIFGYKQRGPAEEALNVFYCYTEGAVDLDHVTDERERKALEGIISNFGQTPCQLLKEP
HPTRLSAEEAAHRLARLDTNSPSIFQHLDELKAFFAEVTVSASGLLGTHSWLPYDRNISNYF
SFSKDPTMGSHKTQRLLSGPWVPGSGVSGQALAVAPDGKLLFSGGHWGSLRVTALPRGKLL
SQLSCHLDVVTCLALDTCGIY LISGSRDTCMVWRL LHQGGLSVGLAPKPVQVLYGHGA AVS
CVAISTELDMAVSGSEDGTVIIHTVRRGQFVAALRPLGATFPGPIFHLALGSEGQIVVQSSA
WERPGAQVTYSLHLYSVNGKLRASLPLAEQPTALTVTEDFVLLGTAQCALHILQLNTLLPAA
PPLPMKVAIRSVAVTKERSHVLVGLEDGKLIVVVAGQPSEVRSSQFARKLWRSSRRISQVSS
GETEYNPTEAR

N-glycosylation site.

amino acids 677-681

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 985-989

Tyrosine kinase phosphorylation site.

amino acids 56-65, 367-376, 543-551

N-myristoylation site.amino acids 61-67, 436-442, 604-610, 610-616, 664-670, 691-697,
706-712, 711-717, 769-775, 785-791, 802-808, 820-826, 834-840,
873-879, 912-918, 954-960

24/330

FIGURE 24

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCC
CACGGCCCACCTTGTGAACTCCTCGTGCCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCAT
CCAAAGGCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTTGGGGCTCTTC
TGGACCCTTAACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTT
CTACTGGGCCTTCCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCC
GCACACTCCGTTACCACACTGGGTCATTGGCATTGAGGCCCTCATCCTGACCCTTGTGCAG
ATAGCCCGGGTCATCTTGGAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGC
CCGCTGCATCATGTGCTGTTTCAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCC
TAAACCGCAATGCATACATCATGATCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAA
AATGCGTTCATGCTACTCATGCGAAACATTGTCAGGGTGGTCGTCCTGGACAAAGTCACAGA
CCTGCTGCTGTTCTTTGGGAAGCTGCTGGTGGTCGGAGGCGTGGGGGTCCTGTCTTCTTTT
TTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAGACTTTAAGAGCCCCCACCTCAACTATTAC
TGGCTGCCCATCATGACCTCCATCCTGGGGGCCTATGTCATCGCCAGCGGCTTCTTCAGCGT
TTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCTGGAAGACCTGGAGCGGAACAACG
GCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAAAGATTCTGGGCAAGAAGAAC
GAGGCGCCCCCGGACAACAAGAAGAGGAAGAAG**TGA**CAGCTCCGGCCCTGATCCAGGACTGC
ACCCACACCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGTCTCCATTTTGTGGT
AAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACTTTGAGAGGCTG
AGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCTCC
GTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCAGCTAC
TCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTGTCAGTGAGCCGAGA
TCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAACAA
AAAGATTTTATTAAAGATATTTTGTTAACCTC

25/330

FIGURE 25

RTRGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLF
WTLNWLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQ
IARVILEYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAK
NAFMLLMRNIVRVVLDKVTDLLLFFGKLLVGGVGVLSFFFFSGRIPGLGKDFKSPHLNYY
WLPIMTSILGAYVIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLLKILGKN
EAPPDNKKRKK

26/330

FIGURE 26

GAGTCTTGACCGCCGCCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGT
GGCT**ATG**TTCGTGTCCGATTTCCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCC
TTCTCTTCGTGGCCTCGGACGTGGATGCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTTC
CAGTGTGACCACGTGCAATATACGCTGGTTCCAGTTTCTGGGTGGCAAGAACTTGAAACTGC
ATTTCTTGAGCATAAAGAACAGTTTCATTATTTTATTCTCATAAACTGTGGAGCTAATGTAG
ACCTATTGGATATTCTTCAACCTGATGAAGACACTATATTCTTTGTGTGTGACTCCCATAGG
CCAGTCAATGTCGTCAATGTATACAACGATACCCAGATCAAATTACTCATTAACAAGATGA
TGACCTTGAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATT
CAGGAAATGACAGTGATGGGTGAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATA
GTGGAGCAAACCATGCGGAGGAGGCAGCGGCGAGAGTGGGAGGCCCGGAGAAGAGACATCCT
CTTTGACTACGAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGATGTTTGAGCTGG
CTTGATGCTGTCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGAC
CAGTGGGTGCAAGACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGTCTTGCAGCG
CCACGTTTCCCGCCACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCA
CACGGATCTCCTTTGAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGAC
AGCCTGTGCAACACCAGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAA
GCGGCTCCAGGAGTTCCTTGACAGACATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAGTTCC
AGGCCATGGACATCTCCTTGAAGGAGAATTTGCGGGAAATGATTGAAGAGTCTGCAAATAAA
TTTGGGATGAAGGACATGCGCGTGACAGACTTTCAGCATTCAATTTGGGTTCAGCACAAAGTT
TCTGGCCAGCGACGTGGTCTTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCT
CAGGGACAGATCACTTCATCCAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTG
TACCATGGCCTGGAACTCGCCAAGAAGCAGCTGCGAGCCACCCAGCAGACCATTGCCAGCTGC
CTTTGCACCAACCTCGTCATCTCCAGGGGCCTTTCCTGTACTGCTCTCTCATGGAGGGCAC
TCCAGATGTCATGCTGTTCTCTAGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCA
AGTCCTTTGTGTGTTTCGACAAAGAACCGGCGCTGCAAACCTGCTGCCCTGGTGATGGCTGCC
CCCCTGAGCATGGAGCATGGCACAGTGACCGTGGTGGGCATCCCCCAGAGACCGACAGCTC
GGACAGGAAGAACTTTTTTGGGAGGGCGTTTGAGAAGGCAGCGGAAAGCACCAGCTCCCGGA
TGCTGCACAACCATTTTGACCTCTCAGTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTT
CTGGACGCACTTATTTCCCTCCTGTCC**TAG**GAATTTGATTCTTCCAGAATGACCTTCTTATT
TATGTAACCTGGCTTTCATTTAGATTGTAAGTTATGGACATGATTTGAGATGTAGAAGCCATT
TTTTATTAAATAAAATGCTTATTTTAGGAAA

27/330

FIGURE 27

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELETAF
LEHKEQFHYFILINCGANVDLLDILQPDEDTIFFVCDSHRPVNVVNVYNDTQIKLLIKQDDD
LEVPAVEDIFRDEEEDEEHSGNDSGSEPSEKRTRLEEEIVEQTMRRRQRREWEARRRDILF
DYEQYEHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGVLQRH
VSRHNHRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSL CNTSYTAARFKLWSVHGQKR
LQEF LADMGLPLKQVKQKFQAMD ISLKENLREMIEESANKFGMKDMRVQTF SIHFGFKHKFL
ASDVVFATMSLMESPEKDGSGTDHFIQALDSLRSNLDKLYHGLELAKKQLRATQQTIASCL
CTNLVISQGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKSFVCSTKNRRCKLLPLVMAAP
LSMEHGTVTTVVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNFHFDLSVIELKAEDRSKFL
DALISLLS

28/330

FIGURE 28

GTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCCGTGGCTATGNTCGTGTCCGATTTCCGCA
AAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCCTTCTCTTCGTGGCCTCGGANGTGGAT
GCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTCCAGTGTGACCANGTGCAATATANGCT
GGTTCCAGTTTCTGGGTGGCAAGAACTTGAACTGCATTTCTTGAGCATAAAGAACAGTTTC
ATTATTTTATTCTCATAAACTGTGGAGCTAATGTAGACCTATTGGATATTCTTCAACCTGAT
GAAGACACTATATTCTTTGTGTGTGACACCCATAGGCCAGTCAATGTTGTCAATGTATACAA
CGATACCC

29/330

FIGURE 29

CAGGAACCCTCTCTTTGGGTCTGGATTGGGACCCCTTTCCAGTACCATTTTTTCTAGTGAAC
CACGAAGGGACGATACCAGAAAAACACCCTCAACCCAAAGGAAATAGACTACAGCCCCAATTG
GCTGACTTTTGGCTATAGAAAAAGAAAGGAACGAAAAGAGACAGTTTTTTTTTGGAAAGCTAA
GTCTTCCCTTTATCGAGTCAAGAAACCCCCCTTCTTGAGCTATTTACAGCTTTTAAACAATT
GAGTAAAGTACGCTCCGGTCACCA**ATG**GTGACAGCCGCCCTGGGTCCCGTCTGGGCAGCGCTC
CTGCTCTTTCTCCTGATGTGTGAGATCCGTATGGTGGAGCTCACCTTTGACAGAGCTGTGGC
CAGCGGCTGCCAACGGTGCTGTGACTCTGAGGACCCCTGGATCCTGCCCATGTATCCTCAG
CCTCTTCCCTCCGGCCGCCCCACGCCCTGCCTGAGATCAGACCCTACATTAATATCACCATC
CTGAAGGGTGACAAAGGGGACCCAGGCCCAATGGGCCTGCCAGGGTACATGGGCAGGGAGGG
TCCCCAAGGGGAGCCTGGCCCTCAGGGCAGCAAGGGTGACAAGGGGGAGATGGGCAGCCCCG
GCGCCCCGTGCCAGAAGCGCTTCTTCGCCTTCTCAGTGGGCCGCAAGACGGCCCTGCACAGC
GGCGAGGACTTCCAGACGCTGCTCTTCGAAAGGGTCTTTGTGAACCTTGATGGGTGCTTTGA
CATGGCGACCGGCCAGTTTGCTGCTCCCGTGCCTGGCATCTACTTCTTCAGCCTCAATGTGC
ACAGCTGGAATTACAAGGAGACGTACGTGCACATTATGCATAACCAGAAAGAGGCTGTCTATC
CTGTACGCGCAGCCCAGCGAGCGCAGCATCATGCAGAGCCAGAGTGTGATGCTGGACCTGGC
CTACGGGGACCGCGTCTGGGTGCGGCTCTTCAAGCGCCAGCGCGAGAACGCCATCTACAGCA
ACGACTTCGACACCTACATCACCTTCAGCGGCCACCTCATCAAGGCCGAGGACGAC**TGA**GGG
CCTCTGGGCCACCCTGCCGGCTGGAGAGCTCAGGTGCTGGTCCCGTCCCCTGCAGGGCTCAG
TTTGCACTGCTGTGAAGCAGGAAGGCCAGGGAGGTCCCCGGGGACCTGGCATTTCTGGGGAGA
CCCTGCTTCTATCTTGGCTGCCATCATCCCTCCCAGCCTATTTCTGCTCCTCTCTTCTCTCT
TGGACCTATTTTTAAGAAGCTTGCTAACCTAAATATTCTAGAACTTTCCCAGCCTCGTAGCCC
AGCACTTCTCAAACCTTGGAATGCATGCGAATCACCCGGGGTTTCGTGTTAAATGCAGATTCT
GACTCAGCAGGTCTGAGTGGGTCCAGGATTCTGTGTTTCTCATATGTTCCCTGGGTGATGCTG
ATGGGGTCAGTCTATGAACCACACTGGAGCAACCAGGTTCTAGGACTTCTCAATATTCTAG
TACTTTCTGAACATTCTGGAATCCTCCCCACATTCTAGAATTCTCCCAACATTTTTTTTTTCT
TGAGACAGAGTCTTGCTCTGTTGCCAGGCTAGAGTGCAGTGGTGCATCTCAGTTCAGTGC
AACCTCTGCCTCCCGGGTTCAAGCGATTCTTCTGCCTCAGCCTCCCTAGTGGCTGGGATTAC
AGGCGCCTGCTACCATGCCTGGCTAATTTTTGTATTTTTAGTAGAGATGGGGTTTCACCAT
TTGGCCAGGCTGGTCTTGAACCTCTGACTTCAGGTGACCCACCCGCCTCGGCCTCTCAAAAT
GCTGGGATTACAGGTGTGAGCCACCGTGCCTGGCCAATTCCAACATTCTTAAATTCTCTCAT
CCCTCCAGGGCTCCCCGTGCTATGTTCTCTTTACCCCTTCCCCCTCTTCTCTTGCTCAGGCC
TGCACCACTGCAGCCACCGTTCATTTTATTCAATTCATTAAACACTGAGCACTCACTCTGTGCT
GGGTCCCGGGAAGGGTGAGGGGGTCAAGACACAGGCCCTGCCCTGCCCCTCAGTGAAGTGGCCA
GTCCAGCCCAGGCGGGGAGAGATGTGTACATAGGTTTTAAAGCAGACCCAGAGCTCATGGGG
GCCTGTGTTCTGGGTGTTCAAGGTGCTGCTGGTCCCTCCATTACCCACTGCTCCCCAAGGCTGG
TGGGACGGGGTCCCGGTGGCAGGGGGCAGGTATCTCCTTCCCGTTCCTCATCCACCTGCCAG
TGCTCATCGTTACAGCAAACCCCAGGGGGCCTTGGCCAGGTCAAGGGTTCTGTGAGGAGAGG
ACCCAGGAGTGTGGGGGCATTTGGGGGGTGAAGTGGCCCCCGAAGAATGGAACCCACACCCA
TAGCTCTCCCCACAGCTGATACGGCATCCTGCGAGAAGACCTGCCCTCCTCACTGGGATCCC
CTTCTGCTCCTCCCAGGGCTCTGCCAGGGCCTTGCTCAGTCCCTTCCACCAAAGTCATCT
GAACTTCCGTTTCCCCAGGGCCTCCAGCTGCCCTCAGACACTGATGTCTGTCCCCAGGTGCT
CTCTGCCCCCTCATGCCCCCTCTCACCGGCCAGTGCCCCGACTCTCCAGGCTTTATCAAGGTG
CTAAGGCCCGGGTGGGCAGCTCCTCGTCTCAGAGCCCTCCTCCGGCCTGGTGCTGCCTTTAC
AAACACCTGCAGGAGAAGGGCCACGGAAGCCCCAGGCTTTAGAGCCCTCAGCAGGTCTGGGG
AGCTAGAGCAAAGGAGGGACCTCAGGCCTTCCGTTTTCTTCTTCCAGGGTGGGGTGGCCTGGT
GTTCCCTAGCCTTCCAAACCCAGGTGGCCTGCCCTTCTCCCCAGAGGGAGGCGGCCTCCGC
CCATTGGTGCTCATGCAGACTCTGGGGCTGAGGTGCCCGGGGGGTGATCTCTGGTGCTCAC
AGCCGAGGGAGCCGTGGCTCCATGGCCAGATGACGGAACAGGGTCTGACCAAGTGCCAGGA
AGACCTGTGCTATAAACACCCTGCCTGATCTGCCCTGCCTGACCCGCCACGCCCTGCC
GTCCAGCATGATTAAAGAATGCTGTCTCCTCTTGGAATAAAAAAAAAAAAAA

30/330

FIGURE 30

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSGRPH
ALPEIRPYINITILKGDKGDPGPMGLPGYMGREGPQGEPPQGSKGDKGEMGSPGAPCQKRF
FAFSVGRKTALHSGEDFQTLLFERVFVNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKET
YVHIMHNQKEAVILYAQPSESRIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYIT
FSGHLIKAEDD

Important features:**Signal peptide:**

amino acids 1-20

N-glycosylation site.

amino acids 72-75

Clq domain proteins.

amino acids 144-178, 78-111 and 84-117

31/330

FIGURE 31

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCCTCGGGCCCGACCCGCCAGGAAAGACTG
AGGCCGCGGCCTGCCCCGCCCCGGCTCCCTGCGCCGCCGCCGCTCCCGGGACAGAAG**ATG**TG
CTCCAGGGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGG
GCTGCCCATCCGGCTGCCAGTGCAGCCAGCCACAGACAGTCTTCTGCACTGCCCCGCCAGGGG
ACCACGGTGCCCCGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCAT
CACCATGCTCGACGCAGGCAGCTTTGCCGGCCTGCCGGGCCTGCAGCTCCTGGACCTGTCAC
AGAACCAGATCGCCAGCCTGCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTG
GACCTGACGGCCAACAGGCTGCATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCT
CGAGCGCCTCTACCTGGGCAAGAACCGCATCCGCCACATCCAGCCTGGTGCCTTCGACACGC
TCGACCGCCTCCTGGAGCTCAAGCTGCAGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGC
CTGCCCCGCTGCTGCTGCTGGACCTCAGCCACAACAGCCTCCTGGCCCTGGAGCCCGGCAT
CCTGGACACTGCCAACGTGGAGGCGCTGCGGCTGGCTGGTCTGGGGCTGCAGCAGCTGGACG
AGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACCTGGATGTGTCCGACAACCAGCTGGAG
CGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACGCGCCTGCGGCTGGCCGGCAACAC
CCGCATTGCCCAGCTGCGGCCCCGAGGACCTGGCCGGCCTGGCTGCCCTGCAGGAGCTGGATG
TGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCCTCTTCCCCCGCCTGCGG
CTGCTGGCAGCTGCCCCGAACCCCTTCAACTGCGTGTGCCCCCTGAGCTGGTTTGGCCCCCTG
GGTGC GCGAGAGCCACGTCACACTGGCCAGCCCTGAGGAGACGCGCTGCCACTTCCCGCCCA
AGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCCAGCCACCACC
ACCACAGCCACAGTGCCCCACCACGAGGCCCGTGGTGCGGGAGCCACAGCCTTGTCTTCTAG
CTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCCAGCCCGCCCTCCA
CTGCCCCACCGACTGTAGGGCCTGTCCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTC
AATGGGGGCACATGCCACCTGGGGACACGGCACCACTGGCGTGCTTGTGCCCCGAAGGCTT
CACGGGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCCAGCCCTACACCAGTCA
CGCCGAGGCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGC
GTGGGGCTGCAGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTA
TCGCAACCTATCGGGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTG
AGTACACGGTCACCCAGCTGCGGCCCCAACGCCACTTACTCCGTCTGTGTATGCCTTTGGGG
CCCCGGGCGGGTGCCGGAGGGCGAGGAGGCCTGCGGGGAGGGCCATACACCCCCAGCCGTCCA
CTCCAACCACGCCCCAGTCACCCAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCCG
CCCTGGCCGCGGTGCTCCTGGCCGCGCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGG
GGGCGGGCCATGGCAGCAGCGGCTCAGGACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCT
GGAAGTGGAGGGAGTGAAGGTCCCCTTGAGGCCAGGCCCGAAGGCAACAGAGGGCGGTGGAG
AGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCCACTCATGGGCTTCCAGGGCCTGGCCTC
CAGTCACCCCTCCACGCAAAGCCCTACATC**TAA**GCCAGAGAGAGACAGGGCAGCTGGGGCCG
GGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCCACACCACGTAAGTTCTCAGTCC
CAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCTGGGCCCTGTTCCCTCTGGA
CCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCCCTAACGTCCCCAGAAC
CGAGTGCCCTATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTCCCTGGGCACGGCG
GGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCCACTCCAGGCGGA
CCCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGCGGCTGTG
TGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAAGGAAGATGCTTTA
GGAACATGTTTTGCTTTTTTAAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCTG
GGAAGATGTTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTGTAAAGACAAACGATGATATG
AAGGCCTTTTGTAAAGAAAAAATAAAAGATGAAGTGTGAAA

32/330

FIGURE 32

MCSRVP L L L L L L L L L L A L G P G V Q G C P S G C Q C S Q P Q T V F C T A R Q G T T V P R D V P P D T V G L Y V F E N
G I T M L D A G S F A G L P G L Q L L D L S Q N Q I A S L P S G V F Q P L A N L S N L D L T A N R L H E I T N E T F R G L R
R L E R L Y L G K N R I R H I Q P G A F D T L D R L L E L K L Q D N E L R A L P P L R L P R L L L L D L S H N S L L A L E P
G I L D T A N V E A L R L A G L G L Q Q L D E G L F S R L R N L H D L D V S D N Q L E R V P P V I R G L R G L T R L R L A G
N T R I A Q L R P E D L A G L A A L Q E L D V S N L S L Q A L P G D L S G L F P R L R L L A A A R N P F N C V C P L S W F G
P W V R E S H V T L A S P E E T R C H F P P K N A G R L L L E L D Y A D F G C P A T T T T A T V P T T R P V V R E P T A L S
S S L A P T W L S P T A P A T E A P S P P S T A P P T V G P V P Q P Q D C P P S T C L N G G T C H L G T R H H L A C L C P E
G F T G L Y C E S Q M G Q G T R P S P T P V T P R P P R S L T L G I E P V S P T S L R V G L Q R Y L Q G S S V Q L R S L R L
T Y R N L S G P D K R L V T L R L P A S L A E Y T V T Q L R P N A T Y S V C V M P L G P G R V P E G E E A C G E A H T P P A
V H S N H A P V T Q A R E G N L P L L I A P A L A A V L L A A L A A V G A A Y C V R R G R A M A A A A Q D K G Q V G P G A G
P L E L E G V K V P L E P G P K A T E G G G E A L P S G S E C E V P L M G F P G P G L Q S P L H A K P Y I

33/330

FIGURE 33

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTGAGCTCAT
CTTCATCATTCATATGAGGAAATAAGTGGTAAAATCCTTGGAATACA**ATG**AGACTCATCAG
AAACATTTACATATTTTGTAGTATTGTTATGACAGCAGAGGGTGATGCTCCAGAGCTGCCAG
AAGAAAGGGAACCTGATGACCAACTGCTCCAACATGTCTCTAAGAAAGGTTCCCGCAGACTTG
ACCCAGCCACAACGACACTGGATTTATCCTATAACCTCCTTTTTCAACTCCAGAGTTTCAAG
TTTTCATTTCTGTCTCCAAACTGAGAGTTTTGATTCTATGCCATAACAGAATTCAACAGCTGG
ATCTCAAAACCTTTGAATTCACAAGGAGTTAAGATATTTAGATTTGTCTAATAACAGACTG
AAGAGTGTAACCTTGGTATTTACTGGCAGGTCTCAGGTATTTAGATCTTTCTTTTAATGACTT
TGACACCATGCCTATCTGTGAGGAAGCTGGCAACATGTACACCTGGAAATCCTAGGTTTGA
GTGGGGCAAAAATACAAAATCAGATTTCCAGAAAATTGCTCATCTGCATCTAAATACTGTC
TTCTTAGGATTCAGAACTCTTCCTCATTATGAAGAAGGTAGCCTGCCCATCTTAAACACAAC
AAAAGTGCACATTGTTTTACCAATGGACACAAATTTCTGGGTTCTTTTGGCTGATGGAATCA
AGACTTCAAAAATATTAGAAATGACAAATATAGATGGCAAAAGCCAATTTGTAAGTTATGAA
ATGCAACGAAATCTTAGTTTGAAGAAATGCTAAGACATCGGTTCTATTGCTTAATAAAGTTGA
TTTACTCTGGGACGACCTTTTCTTATCTTACAATTTGTTTGGCATAACATCAGTGGAACT
TTCAGATCCGAAATGTGACTTTTGGTGGTAAGGCTTATCTTGACCACAATTCATTTGACTAC
TCAATACTGTAATGAGAAGCTATAAAATTTGGAGCATGTACATTTTCAAGAGTGTTTTACATTCA
ACAGGATAAAATCTATTTGCTTTTGACCAAAATGGACATAGAAAACCTGACAATATCAAATG
CACAAATGCCACACATGCTTTTCCCGAATTATCTTACGAAATTTCCAATATTTAAATTTTGCC
AATAATATCTTAACAGACGAGTTGTTTAAAGAACTATCCAAGTGCCTCACTTGAAAACCTCT
CATTTTGAATGGCAATAAACTGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAACACAC
CCTTGGAACACTTGGATCTGAGTCAAAATCTATTACAACATAAAAATGATGAAAATTTGCTCA
TGGCCAGAACTGTGGTCAATATGAATCTGTCTACATAAAATTTGCTGATTCTGTCTTCAG
GTGCTTGCCCAAGTATTCAAATACCTTGACCTAAATAAATAACCAATCCAACTGTACCTA
AAGAGACTATTCATCTGATGGCCTTACGAGAAGCTAAATATTGCATTTAATTTTCTAACTGAT
CTCCCTGGATGCAGTCATTTCAAGTAGACTTTTCAAGTCTGAACATTGAAATGAACCTTCATTCT
CAGCCCATCTCTGGATTTTGTTCAGAGCTGCCAGGAAGTTAAACTCTAAATGCGGGGAAGAA
ATCCATTCGGGTGTACCTGTGAATTAATAAATTTTCAATTCAGCTTGAAACATATTCAGAGGTC
ATGATGGTTGGATGGTCAGATTTCACACCTGTGAATACCTTTTAAACCTAAGGGGAAGTAG
GTTAAAGACGTTTCTCTCCACGAATTATCTTGCAACACAGCTCTGTTGATTGTACCATTTG
TGGTTATTATGCTAGTTCTGGGGTTGGCTGTGGCCTTCTGCTGTCTCCACTTTGATCTGCCC
TGGTATCTCAGGATGCTAGGTCAATGCACACAAACATGGCACAGGGTTAGGAAAACAACCCA
AGAACAACCTCAAGAGAAATGTCCGATTCCACGCAATTTATTTTATACAGTGAACATGATTTCTC
TGTGGGTGAAGAAATGAATTGATCCCCAATCTAGAGAAGGAAGATGGTTCTATCTTGATTTGC
CTTTATGAAAGCTACTTTGACCCTGGCAAAAGCATTAGTGAAAATATTGTAAGCTTCATTGA
GAAAAGCTATAAGTCCATCTTTGTTTTGTCTCCCAACTTTGTCCAGAATGAGTGGTGCCATT
ATGAATTCTACTTTGCCCACCACAATCTCTTCCATGAAAATTTCTGATCATATAATTCTTATC
TTACTGGAACCCATTCCATTCTATTGCATTCCCACCAGGTATCATAAACTGAAAGCTCTCCT
GGAAAAAAAAGCATACTTGGAAATGGCCCAAGGATAGGCGTAAATGTGGGCTTTTCTGGGC
ACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCAGAGAAATGTATGAACTGCAGACA
TTCACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCTCTCTGATGAGAACAGATTGTCT
ATAAAATCCCACAGTCCTTGGGAAGTTGGGGACCACATACTGTTGGGATGTACATTGATA
CAACCTTTATGATGGCAATTTGACAATATTTATTAATAAATAAATAAATGGTTATTCCTTCATA
TCAGTTTCTAGAAGGATTTCTAAGAATGTATCCTATAGAAACACCTTCACAAGTTTATAAGG
GCTTATGGAAAAAGGTGTTTCATCCCAGGATTGTTTATAATCATGAAAAATGTGGCCAGGTGC
AGTGGCTCACTCTTGTAATCCCAGCACTATGGGAGGCCAAGGTGGGTGACCCACGAGGTCAA
GAGATGGAGACCATCCTGGCCAACATGGTGAAACCCTGTCTCTACTAAAAATACAAAATTA
GCTGGGCTGATGGTGCACGCCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGATCG
CTTGAACCCGGGAGGTGGCAGTTGCAGTGAGCTGAGATCGAGCCACTGCACTCCAGCCTGGT
GACAGAGCGAGACTCCATCTCAAAAAAAGAAAAAAGAAAAAATGGAAAACATCC
TCATGGCCACAAAATAAGGTCTAATTCAATAAATTATAGTACATTAATGTAATATAATATTA
CATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCTGGTATGGAAAAACATATTAATAT
GTTATAACTATTAGGTTGGTGCAAACTAATTGTGGTTTTTGCCATTGAAATGGCATTGAA
ATAAAAGTGTAAGAAATCTATACAGATGTAGTAACAGTGGTTGGGCTGGGAGGTTGGA
TTACAGGGAGCATTTGATTTCTATGTTGTGTATTTCTATAATGTTTGAATTTGTTTAGAATGA
ATCTGTATTTCTTTTATAAGTAGAAAAAATAAAGATAGTTTTTACAGCCT

34/330

FIGURE 34

MRLIRNIYIFCSIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQ
LQSSDFHSVSKLRVLILCHNRIQQDLKTFEFNKELRYLDLSNNRLKSVTWYLLAGLRYLDL
SFNDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLP
ILNTTKLHIVLPMDTNFWVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLL
LNKVDLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFR
VFYIQQDKIYLLLTkMDIENLTISNAQMPHMLFPNYPTKFQYLNfANNILTDELfKRTIQLP
HLKTLILNGNKLETLSLVSCFANNTPLEHLDLsqNLLQHkNDENCsWPETVVMNLSYNKLS
DSVFRCLPKSIQILDlnnnQIQTVPKETIHLMAReLNIaFNFLTDLPgCSHFsrLSVLNIE
MNFILSPSLDFVQSCQEVKTLNAGRNPFRCTCELKNFIQLETYSEVMVGWSDSYTCEYPLN
LRGTRLKDVHLHELSCNTALLIVTIVVIMLVGLAVAFcCLHFDLPWYLRMLGQCTQTWHRV
RKTTQEQlKRNVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENI
VSFIEKSYKSI FVLSPNFVQNEWCHYEFYFAHhNLFHENSdHIILILLEPIPFYCIPTRYHK
LKALLEKKAYLEWPKDRRKCGLFWANLRAAINVNVLATREMYELQTFTELNEESRGSTISLM
RTDCL

35/330

FIGURE 35

GGGGGCTTTCTTGGGCTTGGCTGCTTGGAAACACCTGCCTCCAAGGACCGGCCTCGGAGGGGTGCGCGGGAAAGG
GAGGGAAGAAGGAAGGGCGGGGCGGCCCCCTGCGCCCGCCCGCGCCTCTGCGCGCCCCCTGTCCGCCCCGGC
CCAGCCCAGCCCAGCCCCGCGGGCCGGTCACACGCGCAGCCAGCCGCGCCCTCCCGCGCCCAAGCGCGCCGCT
CTGCTGTGCCCTGCGCCCTTGCCCCGCGCCAGCTTCTGCGCCCGCAGCCCCCGCGCCCCCGGTGACCGTGA
CCCTGCCCTGGGCGCGGGGCGGAGCAGGCATGTCCCGCCCGGGGACCGCTACCCAGCGCTGGCCCTGGTGCTC
CTGGCAGTGACCTGGCCGGGGTCCGAGCCCAGGGCGCAGCCCTCGAGGACCCTGATTATTACGGGCAGGAGAT
CTGGAGCCGGGAGCCCTACTACGCGCGCCCGGAGCCCCGAGCTCGAGACCTTCTCTCCGCCGCTGCCTGCGGGGC
CCGGGGAGGAGTGGGAGCGGCGCCCGCAGGAGCCCAGGCCGCCCAAGAGGGCCACCAAGCCCAAGAAAGCTCCC
AAGAGGGAGAAGTCCGGCTCCGGAGCCGCTCCACCAGGTAACACAGCAACAAAAAGTTATGAGAACCAAGAG
CTCTGAGAAGGCTGCCAACGATGATCACAGTGTCCGTGTGGCCCGTGAAGATGTCAGAGAGAGTTGCCACCTC
TTGGTCTGGAAACCTTAAAAATCACAGACTTCCAGCTCCATGCCTCCACGGTGAAGCGCTATGGCCTGGGGCA
CATCGAGGGAGACTCAACATCCAGGCGGGCATTAAATGAAAATGATTTTTATGACGGAGCGTGGTGCGCGGGAAG
AAATGACCTCCAGCAGTGGATTGAAGTGGATGCTCGGCGCCTGACCAGATTCACTGGTGTCACTCAAGGGA
GGAACCTCCCTCTGGCTGAGTGACTGGGTGACATCCTATAAGGTCATGGTGAGCAATGACAGCCACACGTGGGT
ACTGTTAAGAATGGATCTGGAGACATGATATTTGAGGGAACAGTGAGAAGGAGATCCCTGTTCTCAATGAGCT
ACCCGTCCCCATGGTGGCCCGCTACATCCGCATAAACCCCTCAGTCTGGTTTGATAATGGGAGCATCTGCATGA
GAATGGAGATCCTGGGCTGCCACTGCCAGATCCTAATAATTATTATCACCGCCGGAACGAGATGACCACCT
GATGACCTGGATTTTAAGCACCACAATTATAAGGAAATGCGCCAGTTGATGAAAGTTGTGAATGAAATGTGTCC
CAATATCACCAAGATTTACAACATTGGAAAAAGCCACCAGGGCCTGAAGCTGTATGCTGTGGAGATCTCAGATC
ACCTGGGGAGCATGAAGTCGGTGAGCCCCGAGTTCCTACTACATCGCGGGGGGCCACGGCAATGAGGTGCTGGGC
CGGGAGCTGCTGCTGCTGCTGGTGCAGTTCGTGTGTGTCAGGAGTACTTGGCCCGGAATGCGCGCATCTGCACCT
GGTGGAGGAGACGCGGATTACAGTCTCCCTCCCTCAACCCCGATGGCTACGAGAAGGCCTACGAAGGGGGCT
CGGAGCTGGGAGGCTGGTCCCTGGGACGCTGGACCCACGATGGAATTGACATCAACAACAACCTTTCTGATTTA
AACACGCTGCTCTGGGAGGCAGAGGATCGACAGAATGTCCCCAGGAAAGTTCCCAATCACTATATTGCAATCCC
TGAGTGGTTTTCTGTGCGAAAATGCCACGGTGGCTGCCGAGACCAGAGCAGTCATAGCCTGGATGGAAAAATCC
CTTTTGTGCTGGGCGGCAACCTGCAGGGCGGCGAGCTGGTGGTGGCGTATCCCTACGACCTGGTGCAGTCCCC
TGGAAGACGCAGGAACACACCCCCACCCCGATGACCACGTGTTCCGCTGGCTGGCTTACTCCTATGCCTCCAC
ACACCGCCTCATGACAGACGCGCGGAGGAGGGTGTGCCACACGGAGGACTTCCAGAAGGAGGAGGGCACTGTCA
ATGGGGCCTCCTGGCACACCGTCGCTGGAAGTCTGAACGATTTAGCTACCTTCATACAAACTGCTTCGAACTG
TCCATCTACGTGGGCTGTGATAAATACCCACATGAGAGCCAGCTGCCCCAGGAGTGGGAGAATAACCGGGAATC
TCTGATCGTGTTCATGGAGCAGGTTTCATCGTGGCATTAAAGGCTTGGTGAGAGATTACATGGAAAAGGAATCC
CAAACGCCATTATCTCCGTAGAAGGCATTAACCATGACATCCGAACAGCCAACGATGGGGATTACTGGCGCCTC
CTGAACCTGGAGAGTATGTGGTCACAGCAAAGGCCGAAGTTTTACTGCATCCACCAAGAACTGTATGGTTGG
CTATGACATGGGGGCCACAAGGTGTGACTTCACACTTAGCAAAACCAACATGGCCAGGATCCGAGAGATCATGG
AGAAGTTTGGGAAGCAGCCCGTCAGCCTGCCAGCCAGGCGGCTGAAGCTGCGGGGGCGGAAGAGACGACAGCGT
GGGTGACCCCTCCTGGGCCCTTGAGACTCGTCTGGGACCCATGCAATTAAACCAACCTGGTAGTAGCTCCATAG
TGGACTCACTCACTGTTGTTTCTCTGTAATTCAGAAGTGCCTGGAAGAGAGGGTGCATTGTGAGGCAGGTCC
CAAAGGGAAGGCTGGAGGCTGAGGCTGTTTTCTTTCTTTGTTCCCATTTATCCAAATAACTTGGACAGAGCA
GCAGAGAAAAGCTGATGGGAGTGAGAGAACTCAGCAAGCCAACCTGGGAATCAGAGAGAGAAGGAGAAGGAGGG
GAGCCTGTCCGTTCAAGACCTCTGGCTGCATAGAAAAGGATTCTGGTGCTTCCCCTGTTGCGTGGCAGCAAGG
GTTCCACGTGCATTTGCAATTTGCACAGCTAAAATTGCAGCATTTCCCCAGCTGGGCTGTCCCAATGTTACCA
TTTGAGATGCTCCAGGCGTCCTAAGAGAATCCACCCTCTCTGGCCCTGGGACATTGCAAGCTGCTACAAATAA
ATTCTGTGTTCTTTTGACAATAGCGTCATTGCCAAGTGCACATCAGTGAGCCTCTTGAATCTGTTTAGTCTCCT
TTTTCAACAAAGGAGTGTGTTTCAAGAAAAGGAGAGAGAGGCTGAGATCATTAGGAGTTTGTGGGCAGCAAGCA
TGGAGCTTCTTGACAAATTCTGGGTCCATAACAACCCCAAGTCCCTGCTGATCCAGTAGCCCTGGAGGTT
CCCCAGGTAGGGAGAGCCAGAGGTGCCAGCCTTCTGAAGGGCCAGAAAATTTAGCCTGGATCTCCTCTTTTAC
CTGCTAGGACTGGAAGAGCCAGAAGTGGGGTGGCCTGAAGCCCTCTCTGCTTGGGTATTGCCCTGTGTG
GAATTGAGTGCTCATGGGTTGGCCTCATATCAGCCTGGGAGTTATTTTTGATATGTAGAATGCCAGATCTTCCA
GATTAGGCTAAATGTAATGAAAACCTCTTAGGATTATCTGTGGAGCATCAGTTTGGGAAGAATTATTGAATTAT
CTTGCAAGAAAAAGTATGTCTCACTTTTTGTTAATGTTGCTGCCTCATTGACCTGGGAAAAATGAAAAAAA
AATAAAGCAAATGGTAAGACCCTTAAAAAAA

36/330

FIGURE 36

MSRPGTATPALALVLLAVTLAGVGAQGALEDPDYYGQEIWSREPYARPEPELETFSPLP
AGPGEEWERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHS
VRVAREDVRESCPPLGLETLKITDFQLHASTVKRYGLGAHRGRNLNIQAGINENDFYDGAWCA
GRNDLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSHTWVTVKNGSGDMIF
EGNSEKEIPVLNELPVPMPVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTT
TDDLDFKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEF
HYIAGAHGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNPDGYEKAYEGG
SELGGWSLGRWTHDGIDINNNFPDLNTLLWEAEDRQNVPRKVPNHYYIAIPEWFLSENATVAA
ETRAVIAWMEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSAST
HRLMTDARRRVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHES
QLPEEWENNRESLIVFMEQVHRGIKGLVRDSHGKGIPNAIISVEGINHDIRTANDGDYWRL
NPGEYVVTAKAEGFTASTKNCMVGYDMGATRCDFTLSTNMARIREIMEKFGKQPVSLPARR
LKLGRGRRRQRG

37/330

FIGURE 37

CTAAGAGGACAAG**ATG**AGGCCCGGCCTCTCATTCTCCTAGCCCTTCTGTTCTTCCTTGGCCAAGCTGCAGGGG
ATTTGGGGGATGTGGGACCTCCAATTCCCAGCCCCGGCTTCAGCTCTTTCCCAGGTGTTGACTCCAGCTCCAGC
TTCAGCTCCAGCTCCAGGTCCGGGCTCCAGCTCCAGCCGAGCTTAGGCAGCGGAGGTTCTGTGTCCAGTTGTT
TTCCAATTTTACC GGCTCCGTGGATGACCGTGGGACCTGCCAGTGCTCTGTTTCCCTGCCAGACACCACCTTTC
CCGTGGACAGAGTGGAACGCTTGAATTCACAGCTCATGTTCTTTCTCAGAAAGTTTGAGAAAGAACTTTCTAAA
GTGAGGGAATATGTCCAATTAATTAGTGTGTATGAAAAGAACTGTTAAACCTAACTGTCCGAATTGACATCAT
GGAGAAGGATACCATTTCTTACACTGAACTGGACTTCGAGCTGATCAAGGTAGAAGTGAAGGAGATGGAAAAAC
TGGTCATACAGCTGAAGGAGAGTTTTGGTGGAAAGCTCAGAAATTGTTGACCAGCTGGAGGTGGAGATAAGAAAT
ATGACTCTCTTGGTAGAGAAGCTTGAGACACTAGACAAAAACAATGTCCTTGCCATTGCGCGAGAAATCGTGCG
TCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAAACACCCCTGTCGTCCACCCTCCTCCCACTC
CAGGGAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGTGGTTCAGCTCAACTGGAGAGGGTTT
TCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTCCCCAGCATCCAAACAAAGGACTGTATTGGGTGGCGCC
ATTGAATACAGATGGGAGACTGTTGGAGTATTATAGACTGTACAACACACTGGATGATTTGCTATTGTATATAA
ATGCTCGAGAGTTGCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAAC
ATGTACAACACCGGAATATTGCCAGAGTTAACCTGACCACCAACACGATTGCTGTGACTCAAACCTCTCCCTAA
TGCTGCCTATAATAACCGCTTTTTCATATGCTAATGTTGCTTGGCAAGATATTGACTTTGCTGTGGATGAGAATG
GATTGTGGGTTATTTATTCAACTGAAGCCAGCACTGGTAACATGGTGATTAGTAAACTCAATGACACCACACTT
CAGGTGCTAAACACTTGGTATACCAAGCAGTATAAACCATCTGCTTCTAACGCCTTCATGGTATGTGGGGTTCT
GTATGCCACCCGTACTATGAACACCAGAACAGAAGAGATTTTTTACTATTATGACACAAACACAGGGAAAGAGG
GCAAACCTAGACATTGTAATGCATAAGATGCAGGAAAAAGTGCAGAGCATTAACCTATAACCCTTTTGACCAGAAA
CTTTATGTCTATAACGATGGTTACCTTCTGAATTATGATCTTTCTGTCTTGCAAGAGCCCG**TAAG**CTGTTTA
GGAGTTAGGGTGAAAGAGAAAAATGTTTGTGAAAAAATAGTCTTCTCCACTTACTTAGATATCTGCAGGGGTGT
CTAAAAGTGTGTTCAATTTTGCAGCAATGTTTAGGTGCATAGTTCTACCACACTAGAGATCTAGGACATTTGTCT
TGATTTGGTGAGTTCTCTTGGGAATCATCTGCCTCTTCAGGCGCATTTTGCAATAAAGTCTGTCTAGGGTGGGA
TTGTCAGAGGTCTAGGGGCACTGTGGGCCTAGTGAAGCCTACTGTGAGGAGGCTTCACTAGAAGCCTTAAATTA
GGAATTAAGGAACCTTAAACTCAGTATGGCGTCTAGGGATTCTTTGTACAGGAAATATTGCCCAATGACTAGTC
CTCATCCATGTAGCACCCTAATTCTTCCATGCCTGGAAGAACTTGGGGACTTAGTTAGGTAGATTAATATCT
GGAGCTCCTCGAGGGACCAAATCTCCAACCTTTTTTTTCCCCTCACTAGCACCTGGAATGATGCTTTGTATGTGG
CAGATAAGTAAATTTGGCATGCTTATATATTCTACATCTGTAAAGTGCTGAGTTTTATGGAGAGAGGCCTTTTTT
ATGCATTAAATTGTACATGGCAAATAAATCCAGAAGGATCTGTAGATGAGGCACCTGCTTTTTCTTTCTCTC
ATTGTCCACCTTACTAAAAGTCAGTAGAATCTTCTACCTCATAACTTCCTTCCAAAGGCAGCTCAGAAGATTAG
AACCAGACTTACTAACCAATTCCACCCCCCAACCCCTTCTACTGCCTACTTTAAAAAAATTAATAGTTTTT
CTATGGAACCTGATCTAAGATTAGAAAAATTAATTTTCTTAAATTTTATTATGGACTTTTATTTACATGACTCTA
AGACTATAAGAAAAATCTGATGGCAGTGACAAAGTGCTAGCATTATTTGTTATCTAATAAAGACCTTGGAGCATA
TGTGCAACTTATGAGTGTATCAGTTGTTGCATGTAATTTTTTGCCTTTGTTTAAAGCCTGGAACCTGTAAGAAAT
GAAAATTTAATTTTTTTTTTCTAGGACGAGCTATAGAAAAGCTATTGAGAGTATCTAGTTAATCAGTGCAGTAGT
TGGAAACCTTGCTGGTGTATGTGATGTGCTTCTGTGCTTTTGAATGACTTTATCATCTAGTCTTTGTCTATTTT
TCCTTTGATGTTCAAGTCCTAGTCTATAGGATTGGCAGTTTAAATGCTTTACTCCCCCTTTTAAATAAATGAT
TAAAATGTGCTTTGAAAAAATAAAAAAAAAAAAAAAAAAAAAA

38/330

FIGURE 38

MRPGLSFLLALLFFLGQAAGDLGDVGPPIPSPGFSSFPGVDSSSSFSSSSRSGSSSSRSLGS
GGSVSQLFSNFTGSVDDRGTCQCSVSLPDTTFPVDRVERLEFTAHVLSQKFEKELSKVREYV
QLISVYEKKLLNLTVRIDIMEKDTISYTELDFELIKVEVKEMEKLVIQLKESFGGSSEIVDQ
LEVEIRNMTLLVEKLETLDKNNVLAIRREIVALKTKLKECEASKDQNTPVVHPPPTPGSCGH
GGVVNISKPSVVQLNWRGFSYLYGAWGRDYSPQHHPNKGLYWVAPLNTDGRLLLEYRLYNTLD
DLLLYINARELRITYGQGSGTAVYNNNMYVNMYNTGNIARVNLTNTIAVTQTLPNAAYNNR
FSYANVAWQDIDFAVDENGLWVIYSTEASTGNMVISKLNDDTLQVLNTWYTKQYKPSASNAF
MVCGLVLYATRTMNTRTEEIFYYYDTNTGKEGKLDIVMHKMQEKVQSINYNPFDQKLYVYNDG
YLLNYDLSVLQKPQ

39/330

FIGURE 39

GCTCTGAAGACCAAGCTGAAAGAGTGTGAGGCCTCTAAAGATCAAACACCCCTGTCGTCCAC
CCTCCTCCCCTCCAGGGAGCTGTGGTCATGGTGGTGTGGTGAACATCAGCAAACCGTCTGT
GGTTCAGCTCAACTGGAGAGGGTTTTCTTATCTATATGGTGCTTGGGGTAGGGATTACTCTC
CCCAGCATCCAAACAAAGGNATGTATTGGGNGGCGCCATTGAATACAGATGGGAGACTGTTG
GAGTATTATAGACTGTACAACCCACTGGATGATTTGCTATTGTATATAAATGCTCGAGAGTT
GCGGATCACCTATGGCCAAGGTAGTGGTACAGCAGTTTACAACAACAACATGTACGTCAACA
TGTACAACACCGGNATATTGCCAGAGTTAACCTGACC

40/330

FIGURE 40

TCTCGCAGATAGTAAATAATCTCGGAAAGGCGAGAAAGAAGCTGTCTCCATCTTGTCTGTAT
CCGCTGCTCTTGTGACGTTGTGGAG**ATG**GGGAGCGTCCTGGGGCTGTGCTCCATGGCGAGCT
GGATACCATGTTTGTGTGGAAGTGCCCCGTGTTTGCTATGCCGATGCTGTCTAGTGGAAAC
AACTCCACTGTAAC TAGATTGATCTATGCACTTTTCTTGCTTGTGGAGTATGTGTAGCTTG
TGTAATGTTGATACCAGGAATGGAAGAACAACCTGAATAAGATTCCTGGATTTTGTGAGAATG
AGAAAGGTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTT
GGTTTGGCTATGTTCTATCTTCTTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGA
TCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTA
TTATTGGGGCATTCTTCATTCCAGAAGGAACCTTTACAACCTGTGTGGTTTTATGTAGGCATG
GCAGGTGCCTTTTGTTCATCCTCATAACAACCTAGTCTTACTTATTGATTTTGCACATTCATG
GAATGAATCGTGGGTTGAAAAAATGGAAGAAGGGAACCTCGAGATGTTGGTATGCAGCCTTGT
TATCAGCTACAGCTCTGAATTATCTGCTGTCTTTAGTTGCTATCGTCCTGTTCTTTGTCTAC
TACACTCATCCAGCCAGTTGTTTCAAGAAACAGGCGTTTATCAGTGTCAACATGCTCCTCTG
CGTTGGTGCTTCTGTAATGTCTATACTGCCAAAAATCCAAGAATCACAACCAAGATCTGGTT
TGTTACAGTCTTCAGTAATTACAGTCTACACAATGTATTTGACATGGTCAGCTATGACCAAT
GAACCAGAAACAAATTGCAACCCAAGTCTACTAAGCATAATTGGCTACAATACAACAAGCAC
TGTCCTCAAAGGAAGGGCAGTCAGTCCAGTGGTGGCATGCTCAAGGAATTATAGGACTAATTC
TCTTTTGTGTGTGTATTTTATTCCAGCATCCGTAAGTCTCAAACAATAGTCAGGTTAATAAA
CTGACTCTAACAAAGTGATGAATCTACATTAATAGAAGATGGTGGAGCTAGAAGTGATGGATC
ACTGGAGGATGGGGACGATGTTTCAACGAGCTGTAGATAATGAAAGGGATGGTGTCACTTACA
GTTATTCCTTCTTTCACCTTCATGCTTTTCTGGCTTCACTTTATATCATGATGACCCTTACC
AACTGGTCCAGGTATGAACCTCTCGTGAGATGAAAAGTCAGTGGACAGCTGTCTGGGTGAA
AATCTCTTCCAGTTGGATTGGCATCGTGCTGTATGTTTGGACACTCGTGGCACCCTTGTTC
TTACAAATCGTGATTTT**GAC****TGAG**TGAGACTTCTAGCATGAAAGTCCCCTTTGATTATTGC
TTATTTGAAAACAGTATTCCCAACTTTTGTAAAGTTGTGTATGTTTTTGTCTCCCATGTAAC
TTCTCCAGTGTTCTGGCATGAATTAGATTTTACTGCTTGTCAATTTTGTATTTTCTTACCAA
GTGCATTGATATGTGAAGTAGAATGAATTGCAGAGGAAAGTTTTATGAATATGGTGATGAGT
TAGTAAAAGTGGCCATTATTGGGCTTATTCTCTGCTCTATAGTTGTGAAATGAAGAGTAAAA
ACAAATTTGTTTGACTATTTTAAAATTATATTAGACCTTAAGCTGTTTTAGCAAGCATTAAA
GCAAATGTATGGCTGCCTTTTGAATATTTGATGTGTTGCCTGGCAGGATACTGCAAAGAAC
ATGGTTTTATTTTAAAATTTATAACAAGTCACTTAAATGCCAGTTGTCTGAAAAATCTTATA
AGGTTTTACCCTTGATACGGAATTTACACAGGTAGGGAGTGTTTAGTGGACAATAGTGTAGG
TTATGGATGGAGGTGTCCGTACTAAATTGAATAACGAGTAAATAATCTTACTTGGGTAGAGA
TGGCCTTTGCCAACAAAGTGAACCTGTTTTGGTTGTTTTAAACTCATGAAGTATGGGTTTCACT
GGAAATGTTTGAACCTCTGAAGGATTTAGACAAGGTTTTGAAAAGGATAATCATGGGTTAGA
AGGAAGTGTTTTGAAAGTCACTTTGAAAGTTAGTTTTGGGCCAGCACGGTAGCTCACCTT
GGTAATCCCAGCACTTTGGGAGCTTAAGTGGGTAGATTACTTGAGCCCAGGAATTCAGACCA
GCTTGGCACATGGTGAACCTGTTCTATAAAAAAATCTGGCTTTGAGCATATGCCTGTGGTC
CAGCACTGAGAGGCTAGTGAAGATTGCTGAGCCCAGAGCCAAAGGTTGCAGTGAGCAAGTCA
CGTCACTGCACTCTAGCTGGCACAGAGTAAGCCAAAAAATATATATATATTGAAATCAAGG
AGGCAAAATTTTACAGGGAAGGAAGTAAGTGCAAAACCACTAGGCTTTAGTAGGTACTTAT
ATAAAATCTAGTCCAGTTCTCTCATTTAAAAAATGAAGACACTGAAATACAGACTTAAATA
GCTCAGATAGCTAATTAGGAAATTTCAAGTTGGCCAATAATAGCATTCTCTCTGACATTTAA
AAATAATTTCTATTCAAAATACATGCATATTGATTTACACCTCATACTGTGATAATTAATGT
GATGTGGATTGCTGGTGTCCAGCATGACCCATAAACAGGTCAGAAGAATGATGGAATGTTTT
AGAATAAACTCCTGCTTATAGTATACTACACAGTTCAAAGATGTTTAAATGCTTTTGTAT
TTACTGCCATGTAATTGAAATATATAGATTATTGTAACCTTTCAACCTGAAAATCAAGCAGT
ATGAGAGTTTAGTTATTTGTATGTGTCTAGTGTCTAATGAAGCTTTTAAATCTACAAAT
TCTTCTTTAAAAATATTTAATGTGAATGGAATATAACAATTCAGCTTAAATCCCCAACCC
TTATTCTGTGTGTAGACATTGTATTCCACAATTTTGAATGGCTGTGTTTTACCTCTAAATAA
ATGAATTCAGAGAAAAA

41/330

FIGURE 41

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGME
EQLNKIPGFCENEKGVVPCNILVGKAVYRLCFGLAMFYLLLSLLMIKVKSSSDPRAAVHNG
FWFFKFAAAIAIIIGAFFIPEGTFTTVWFYVGMAGAFCFILIQVLVLLIDFAHSWNESWVEKM
EEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPASCSSENKAFISVNMLLCVGASVMSI
LPKIQESQPRSGLLQSSVITVYTMyltWSAMTNEPETNCNPSLLSIIGYNTTSTVPKEGQSV
QWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTSDESTLIEDGGARSDGSLEDGDDVH
RAVDNERDGVTYSSFFHFMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVWVKISSSWIGI
VLYVWTLVAPLVLTNRDFD

42/330

FIGURE 42

GCGAGAAAGAAGCTGTCTCCATCTTGTCTGTATCCCGCTGCTTCTTGNGACGTTGTGGAGAT
GGGGAGCGTCCCTGGGGCTGTGCTCCATGGCGAGCTGGATAACCATGTTTGTGTGGAAGTGCC
CCGTGTTTGCTATGCCGATGCTGTCCTAGTGGAAACAANTCCACTGTAAGTAGATTGATCTA
TGCACTTTTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAG
AACAACTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATT
TTGGTTGGCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCT
CTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGAT
TTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

43/330

FIGURE 43

GTTATTGTGAACTTTGTGGAGATGGGAGGTCNTGGGGCTGTGTTCCATGGCGAGCTGGATAC
CANGTTTGTGTGGAAGTGCCCCGTGTTTGNTATGCCGATGCTGTCCTAGTGGAACAANTCC
ACTGTAATTAGATTGATNTATGCACTTTTNTTGCTTGTTGGAGTANGTGTAGCTTGTGTAAT
GTTGATACCAGGAATGGAAGAACAACCTGAATAAGATTCCTGGATTTTGTGAGAATGAGAAAG
GTGTTGTCCCTTGTAACATTTTGGTTGGCTATAAAGCTGTATATNGTTTGTGCTTTGGTTTG
GCTANGTTCTATNTTCTTCTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAG
AGCTGCAGTGCACAATGGATTTTGGTTTTTTAAATTTGCTGCAGCAATTGCAATTATTATTG
GGGC

44/330

FIGURE 44

AAGAAGCTGTCTCCATCTTGTCTGTATCCGCTGCTCTTGTGAACGTTNTGGAGATGGGGAGC
GTCCTTGGGGTTGTGCTCCATGGCGAGCTGGATAACCATGTTTGTGTGGAAGTGCCCCGTGTT
TGCTATGCCGATGCTGTCCTAGTGGAAACAACCTCCACTGTAACTAGATTGATCTATGCACTT
TTCTTGCTTGTTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAGAACAAC
GAATAAGATTCCTGGATTTTGTGAGAATGAGAAAGGTGTTGTCCCTTGTAACATTTTGGTTG
GCTATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTCTTA
CTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTT
CTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGC

45/330

FIGURE 45

GCTGTCCTTAGTGGAACAANTCCAACCTTGTAACCTGGATTGATCTATGCACTTTTTTCCTTG
CTTGTTGGAGTATGTGTAGCTTTGTGTAATGTTGTTCCCAGGATTGGANGAACAACTGAATA
AGATTCCTGGATTTTTGTGAGAATGAGAAAGGTGTTGTCCCCTTGTAACATTTTTGGTTGGC
TATAAAGCTGTATATCGTTTGTGCTTTGGTTTGGCTATGTTCTATCTTCTTCTCTTTACT
AATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTGCACAATGGATTTTGGTTCT
TTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGCATTCTTCATTCCAGAAGGAACCTTT
ACAACTGTGTGGTTTTATGTAGGCATGGCAGGTGCCTTTTGTTCATCCTCATACTAGT
CTTACTTATTGATTTTGCACATTCATGGAATGAATCGTGGGTGAAAAAATGGAAGAAGGGA
ACTCGAGATGTTGGTATGCAGCCTTGTTATCAGCTACAGCTCTGAATTATCTGCTGTCTTTA
GTTGCTATCGTCCTGTTCTTTGTCTACTACACTCATCCAGCCAGTTGTTTCAGAAAACAAGGC
GTTTCATCAGTGTCAACATGCTCCTCTGCGTTGGTGCTTCTGTAATG

FIGURE 46

CTCTGGGCGCGCACAGGCAGTCTCGGTTTTGCCCTGCGATTGAGCTGCGGGTTCGCGGGCGCGCGGCGCTCTCCAAT
GGCAAATGTGTGTGGCTGGAGGCGAGCGCGAGGCTTTTCGGCAAAGGCAGTCGAGTGTTCGAGACCGGGGCGAG
TCCTGTGAAAGCAGATAAAAGAAAACATTTATTAACGTGTCATTACGAGGGGAGCGCCCCGGCCGGGGCTGTCCG
ACTCCCCCGGGAACATTTGGCTCCCTCCAGCTCCGAGAGAGGAGGAAGAAGCGGAAAGAGGCAGATTCA
GTGTTTTCCAGCCAAAGTGGACCTGATCGATGGCCCTCTGAATTTATACAGATATTTGATTTATTAGCATGCCC
CCCTGGTTTTGTGTGTTACGCACACACACAGCTGCACACAAGGCTCTGGCTCGCTTCCCTCCCTCGTTTCCAGCTCC
TGGGCGAATCCCACATCTGTTTCAACTCTCCGCCGAGGGCGAGCAGGAGCGAGAGTGTGTGAATCTGCGAGTG
AAGAGGGACGAGGGAAAAGAAACAAAGCCACAGACGCAACTTGAGACTCCCGCATCCCAAAGAGCACCAGAT
CAGCAAAAAAGAAATGGGGCCCCCGAGCCTCGTGCTGTGCTTGTGTCCGCAACTGTGTTCTCCCTGCTGG
TGGAAGCTCGGCCTTCTGTGCGACACCACCGGCTGAAAGGCAGGTTTCAGAGGGACCGCAAGCAACATCCGCCCA
ACATCATCTGGTGCTGACGGACGACCGAGTGTGGAGCTGGTTCATGCAAGTGATGAACAAGACCCGGCG
ATCATGGAGCAGGGCGGGGCGCACTTCATCAACGCCTTCGTGACCACACCCATGTGCTGCCCTCAGCTCCTC
ATCCTCACTGGCAAGTACGTCCACAACCACAACACCTACACCAACAATGAGAACTGCTCCTCGCCCTCCTGGC
AGGCACAGCACGAGAGCCGCACCTTTGCCGTGTACCTCAATAGCACTGGCTACCGGACAGCTTTCTTCGGGAAG
TATCTTAATGAATACAACGGCTCCTACGTGCCACCCGGCTGGAAGGAGTGGGTGCGACTCCTTAAAAACTCCCG
CTTTTATAACTACACGCTGTGTGCGGAACGGGTGAAAGAGAAGCAGGCTCCGACTACTCCAAGGATTACCTCA
GACACTCATCACCATGACAGCGTGAGCTTCTTCGCAAGTCCAAGAGATGTACCCGACAGGCCAGTCCCTC
ATGGTCACTCAGCCATGCAGCCCCCAGCGCCCTGAGGATTACGCCCAACAATATTACGCCTCTTCCCAAACGC
ATCTCAGCACATCAGCCGAGCTACAACCTACGCGCCCAACCCGGACAAACACTGGATCATGCGCTACACGGGGC
CCATGAAGCCCATCCACATGGAATTCACCAACATGCTCCAGCGGAAGCGCTTGACAGCCCTCATGTGCGTGGAC
GACTCCATGGAGACGATTTACAACATGCTGGTTGAGACGGGCGAGCTGGACAACACGTATAGCTGATATACACCGC
CGACCACGGTTACCACATCGGCCAGTTTGGCCTGGTGAAAGGGAATCCATGCCATATGATTTGACATCAGGG
TCCCGTTCTACGTGAGGGGGCCCCAACGTGGAAGCGCGTGTGTAATCCCCACATCGTCTCAACATTGACCTG
GCCCCACCATCCTGGACATTGCAGGCCTGGACATACCTGCGGATATGGACGGGAATCCATCCTCAAGCTGCT
GGACACGGAGCGGCCCGGTGAATCGGTTTCACTTGAAAAAGAAGATGAGGGTCTGGCGGGACTCCTTCTTGGTGG
AGAGAGGCAAGCTGTACACAAGAGAGACAATGACAAGGTGGACGCCCAGGAGGAGAACTTTCTGCCAAGTAC
CAGCGTGTGAAGGACCTGTGTACGCTGCTGAGTACCAGACGGCGTGTGAGCAGCTGGGACAGAAGTGGCAGTG
TGTGAGGAGCGCCACGGGAAGCTGAAGCTGCATAAGTGCAAGGGCCCCATCGGCTGGGCGGCAGCAGACGCC
TCTCCAACCTCGTGCCCAAGTACTACGGGCAGGGCAGCGAGGCGTGCACCTGTGACAGCGGGGACTACAAGCTC
AGCCTGGCCGACGCGCCGAAAAAACTCTTCAAGAAGAAGTACAAGGCCAGCTATGTCCGCGAGTCGCTCCATCCG
CTCAGTGGCCATCGAGGTGGACGGCAGGGTGTACCACGTAGGCCTGGGTGATGCCGCCAGCCCCGAAACCTCA
CCAAGCGGCACTGGCCAGGGGGCCCTGAGGACCAAGATGACAAGGATGGTGGGGACTTCAGTGGCACTGGAGGC
CTTCCCAGCTACCTCAGCCGCCAACCCCATTAAGTGACACATCGGTGCTACATCCTAGAGAACGACACAGTCAA
GTGTGACCTGGACCTGTACAAGTCCCTCGAGGCTGGAAGACCAAGCTGCACATCGACCAAGAGATTTGAAA
CCCTGCAGAACAAAATTAAGAACCTGAGGGAAGTCCGAGGTCACTGAAGAAAAAGCGGCCAGAAGAATGTGAC
TGTACAAAATCAGCTACCACACCCAGCACAAAGGCCGCCTCAAGCACAGAGGCTCCAGTCTGCATCCTTTCAG
GAAGGGCCTGCAAGAGAAGGACAAGGTGTGGCTGTTGCGGGAGCAGAAGCGCAAGAAGAACTCCGCAAGCTGC
TCAAGCGCCTGCAGAACAACGACACGTGCAGCATGCCAGGCCTCACGTGCTTCACCCACGACAACCAGCACTGG
CAGACGGCGCCTTTCTGGACACTGGGGCCTTTCTGTGCTGCACAGCGCCAAACAATAACACGTACTGGTGCAT
GAGGACCATCAATGAGACTACAATTTCTCTGTGTGAATTTGCAAGCTGGCTTCTTAGAGTACTTTGATCTCA
ACACAGACCCCTACAGCTGATGAATGCAGTGAACACACTGGACAGGGATGTCCTCAACCAGCTACACGTACAG
CTCATGGAGCTGAGGAGCTGCAAGGGTTACAAGCAGTGTAAACCCCGGACTCGAAACATGGACCTGGATGGAGG
AAGCTATGAGCAATACAGGCAGTTTCAGCGTCGAAAGTGGCCAGAAATGAAGAGACCTTCTTCAAATCACTGG
GACAACTGTGGGAAGGCTGGGAAGGTAAAGAAACAACAGAGGTGGACCTCCAAAAACATAGAGGCATCACCTGA
CTGCACAGGCAATGAAAAACCATGTGGGTGATTTCCAGCAGACCTGTGCTATTGGCCAGGAGGCCTGAGAAAGC
AAGCAGCACTCTCAGTCAACATGACAGATTCTGGAGGATAACACGAGGAGCAGAGATAAATTCAGGAAGTCC
ATTTTTGCCCTGCTTTTGCCTTGTGATTATACCTCACCAGCTGCACAAAATGCATTTTTTTCGTATCAAAAAGTC
ACCACTAACCTCCCCCAGAAGCTCACAAAGGAAAACGGAGAGAGCGAGCGAGAGAGATTTCTTGGAAATTTT
TCCCAAGGGCGAAAGTCATTGGAATTTTTAAATCATAGGGGAAAAGCAGTCCTGTTCTAAATCCTCTTATTCTT
TTGGTTTGTACAAAGAAGGAACTAAGAAGCAGGACAGAGGCAACGTGGAGAGGCTGAAAACAGTGCAGAGACG
TTTGACAAATGAGTCAGTAGCAAAAAGAGATGACATTTACCTAGCACTATAAACCTGGTTGCCTCTGAAGAAA
CCCACTTTTTCAGGAGTGGTGGTGTAATAAACGCTCTGTGGCAGTGTAAGAAAAA

47/330

FIGURE 47

MGPPSLVLCLLSATVFSLLGGSSAFLSHHRLKGRFQDRRNIRPNIIILVLTDDQDVELGSMQ
VMNKTRRIMEQGGAHFINAFVTT PMCCPSRSSILT GKYVHNHNTYTNNE NCSSPSWQAQHE S
RTFAVYLNSTGYRTAFFGKYLNEYNGSYVPPGWKEWVGLLKNSRFYNYTL CRNGVKEKHGSD
YSKDYLTDLITNDSVSFFRTSKKMYPHRPVLMVISHAAPHGPEDSAPQYSRLFPNASQHITP
SYNYAPNPDKHWIMRYTGPMKPIHMEFTNMLQRKRLQTLMSVDDSMETIYNMLVETGELDNT
YIVYTADHGYHIGQFGLVKGKSMPYEFDIRVPFYVRGPNVEAGCLNPHIVLNIDLAPTILDI
AGLDIPADM DGKSILKLLDTERPVNRFHLKKKMRVWRDSFLVERGKLLHKRDNDKVDAQEEN
FLPKYQRVKDL CQRAEYQTACEQLGQKWQCVEDATGKLKLHKCKGPMRLGGSRALSNLVPKY
YGQGSEACTCD SGDYKLSLAGRRKKLFKKKYKASYVRSRSIRSV AIEVDGRVYHVGLGDAAQ
PRNLTKRHWP GAPEDQDDKDGGDFSGTGGLPDYSAANPIKVTHRCYI LENDTVQCDLDLYKS
LQAWKDHKLHIDHEIETLQNKIKNLREVRGHLKKKRPEECDCHKISYHTQHKGRLKHRGSSL
HPFRKGLQE KDKVWLLREQKRKKLRKLLKRLQNNDTCSMPGLTCFTHDNQHWQTAPFWTLG
PFCACTSANNNTY WCMRTINETHNFLFCEFATGFLEYFDLNTDPYQLMNAVNTLDRDVLNQL
HVQLMELRSCKGYKQCNPRTRNMDLDGGSYEQYRQFQRRKWPEMKRPSSKSLGQLWEGWEG

48/330

FIGURE 48

AACAAAGTTCAGTGACTGAGAGGGCTGAGCGGAGGCTGCTGAAGGGGAGAAAGGAGTGAGGA
GCTGCTGGGCAGAGAGGGACTGTCCGGCTCCCAG**ATG**CTGGGCCTCCTGGGGAGCACAGCCC
TCGTGGGATGGATCACAGGTGCTGCTGTGGCGGTCCTGCTGCTGCTGCTGCTGCTGGCCACC
TGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGAGGAACCGTACAGCTGCAGGGGGAAA
CCGAGTCCGCCGGGCCCAGCCTTGGCCCTTCCGGCGGCGGGGCCACCTGGGAATCTTTCACC
ATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCAC
CCCCGCCACACCCCTCACCACCTCCACCACCACCACCACCCCCACCGCCACCATCCCCGCCA
CGCTCGC**TGA**GGCTGCTGTGCGCCGGTGCCTGTGGACAGCAGCTGCCCCCTGCCCTCCCATCTG
TCCCAGGACAAGTGGACCCCATGTTTCCATGTGGAAGGATGCATCTCTGGGGTGAACGAGG
GGAACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGCATGCCCCAGTGTACTATGGC
AGCAGAGAATGGAGGAACACTGGGTCTGCAGTGCTGAAGGGTTTGGGGAGTGGAGAGCAAGG
GTGCTCTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGACTCCCAGTGAGCCCCAGAAATG
ACAAGCGTGTCTTGGCAGAGCCAGCACACAAGTGGATGTGAAGTGCCCGTCTTGACCTCCTC
ATCAGGCTGCTGCAGGCCTCTGGCGGGCAGGGCACTGGGAGAGGCCCTGAGAATGTCCTTTT
GGTTTGGAGAAGGCAGTGTGAGGCTGCACAGTCAATTCATCGGTGCCTTAGTCCAAGAAAAT
AAAAACCACTAAGAAGCTTTAAAAAAAAAAAAAAAAAAAAA

49/330

FIGURE 49

MLGLLGSTALVGWITGA AVAVLLLLLLLLLATCLFHGRQDCDVERNRTAAGGNRVRRRAQPWPFRR
RGHLGIFHHHRHPGHVSHVPNVGLHHHHHPRHTPHHLHHHHHPRHHHPRHAR

50/330

FIGURE 50

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTTGGGGATCCAGAGCC**ATG**TCGGACCTGCTA
CTACTGGGCCTGATTGGGGGCCTGACTCTCTTACTGCTGCTGACGCTGCTGGCCTTTGCCGG
GTACTCAGGGCTACTGGCTGGGGTGGAAAGTGAGTGCTGGGTACCCCCCATCCGCAACGTCA
CTGTGGCCTACAAGTTCCACATGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGC
TGCAGCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCC
CCCTGATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCC
CTGAGCTCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGC
CATGTGGTGACAGCCACCTTCCCCTACACCACCATTTCTGTCCATCTGGCTGGCTACCCGCCG
TGTCCATCCTGCCTTGGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGG
AGATCTACCAGGAAGACCAGATCCATTTTCATGTGCCCACTGGCACGGCAGGGAGACTTCTAT
GTGCCTGAGATGAAGGAGACAGAGTGGAATGGCGGGGGCTTGTGGAGGCCATTGACACCCA
GGTGGATGGCACAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCC
CTGGCAGCCGGGAGACTTCAGCTGCCACACTGTCACCTGGGGCGAGCAGCCGTGGCTGGGAT
GACGGTGACACCCGCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGA
GGAGCTGGACTTGGAGGGCGAGGGGCCCTTAGGGGAGTCACGGCTGGACCCTGGGACTGAGC
CCCTGGGGACTACCAAGTGGCTCTGGGAGCCCACTGCCCCTGAGAAGGGCAAGGAG**TAA**CCC
ATGGCCTGCACCCTCCTGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCT
CCAGCCCTCTTCCTCCTTCCTCTGGGGGAGGAGGGGTTCTGAGGGACCTGACTTCCCCTGC
TCCAGGCCTCTTGCTAAGCCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCA
GGGACTATTTTCTGCACCAGCCCCCAGGGCTGCCGCCCTGTTGTGTCTTTTTTTTCACTC
ACAGTGGAGCTTCCAGGACCCAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAA
AAAAA

51/330

FIGURE 51

MSDLLLLGLIGGLTLLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGR
LFTESCSISPKLRSIAYYYDNPHMVPPDKCRAVGSI LSEGEESPSPELIDLYQKFGFKVFS
FPAPSHVVTATFPYTTILSIWLATTRVHPALDTYIKERKLCAYPRLEIYQEDQIHFMCP LAR
QGDFYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSAATLSPGAS
SRGWDDGDTRSEHSYSESGASGSSFEELDLEGEGLGESRLDPGTEPLGTTKWLWEPTAPEK
GKE

52/330

FIGURE 52

CCGCGGGAACGCTGTCCTGGCTGCCGCCACCCGAACAGCCTGTCCTGGTGCCCCGGCTCCCT
GCCCCGCGCCCAGTC**ATG**ACCCTGCGCCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCT
GCTGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGA
CCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGA
GACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCT
GACCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGA
GTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCCTTCTCACTTGGCCTAT
GGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCT
GATTGCACTAATCCGAGCCAACTACTGGCTAAAGCTGGTGAAGGGCATTTTGCCTCTGGTAG
GGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTTGGGTATCACCTATACAGAAAGGCCAAT
AGACCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAA**TA**
ATAAATAATAAATTTTAAAAAACTTAAAAAAAAAAAAAAAAAAAA

53/330

FIGURE 53

MTLRPSLLPLHLLLLLLLLSAAVCRAEAGLETESPVRTLQVETLVEPPEPCAEPAAFGDTLHI
HYTGSLVDGRIIDTSLTRDPLVIELGQKQVIPGLEQSLDMCVGEKRRATIPSHLAYGKRGF
PPSVPADAVVQYDVELIALIRANYWLKLVKGILPLVGMAMVPALLGLIGYHLYRKANRPKVS
KKKLKEEKRNKSKKK

54/330

FIGURE 54

CCCGGGAACGTGTTCCCTGGCTGCCGCACCCGAACAGCCTGTCCTGGTGCCCCGGCTCCCTGC
CCCGCGCCCAGTCATGACCCTGCGCCCCTCACTCCTCCCGCTCCATCTGCTGCTGCTGCTGC
TGCTCAGTGCGGCGGTGTGCCGGGCTGAGGCTGGGCTCGAAACCGAAAGTCCCGTCCGGACC
CTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCCGAGCCCGCTGCTTTTGGAGA
CACGCTTCACATACTACACGGGAAGCTTGGTAGATGGACGTATTATTGACACCTCCCTGA
CCAGAGACCCTCTGGTTATAGAACTTGGCCAAAAGCAGGTGATTCCAGGTCTGGAGCAGAGT
CTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATTCTTCTCACTTGGCCTATGG
AAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGTGCAGTATGACGTGGAGCTGA
TTGCACTAATCCGAGCCAACCTACTGGCTAAAGCTGGTGAAGGGCATTTTGCCTCTGGTAGGG
ATGGCCATGGTGCCACCCTCCTGGGCCTCATTGGGTATCACCTATACAGAAAGGCCAATAGA
CCCAAAGTCTCCAAAAAGAAGCTCAAGGAAGAGAAACGAAACAAGAGCAAAAAGAAATAATA
AATAATAAATTTTAAAAAACTTA

55/330

FIGURE 55

CCGAAAGTCCCGTCCGGACCCTCCAAGTGGAGACCCTGGTGGAGCCCCCAGAACCATGTGCC
GAGCCCGCTGCTTTTGGAGACACGCTTCACATACACTACACGGGAAGCTTGGTAGATGGACG
TATTATTGACACCTCCCTGACCAGAGACCCTCTGGTTATAGAAGCTTGGCCAAAAGCAGGTGA
TTCCAGGTCTGGAGCAGAGTCTTCTCGACATGTGTGTGGGAGAGAAGCGAAGGGCAATCATT
CCTTCTCACTTGGCCTATGGAAAACGGGGATTTCCACCATCTGTCCCAGCGGATGCAGTGGT
GCAGTATGACGTGGAGCTGATTGCACTAATCCGAGCCAACACTGGCTAAAGCTGGTGAAGG
GCATTTTGCCTCTGGTAGGGATGGCCATGGTGCCAGCCCTCCTGGGCCTCATTGGGTATCAC
CTATACAGAAAGGCCAATAGACCCAAAGTCTCCAAAAGAAGCTCAAGGAAGAGAAACGAAA
CAAGAGCAAAAAGAAATAATAAATAATAAATTTTAAAAAACTTAAAA

56/330

FIGURE 56

CTGCTGCATCCGGGTGTCTGGAGGCTGTGGCCGTTTTGTCTTGGCTAAAATCGGGGGAG
TGAGGCGGGGCCGGCGCGGCGGACACCGGGCTCCGGAACCACTGCACGACGGGGCTGGACTG
ACCTGAAAAAA**ATG**CTCTGGATTTCTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGG
GAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGAT
TATCATAGATGCAGCTGTTATTTATCCACCATGAAAGATTTCAACCACTCATACCATGCCT
GTGGTGTATAGCAACCATAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGA
GGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTGCTCGCATTTGGCTTTTCGTTGG
TTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTG
CTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATCTTT
TTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGACTTATGGCAG**TGA**ACACATCTGAT
TTCCCACAGCACAAACAGCCCTGCATGGGTTTGTGTTGTTTTTTACTGCTCACTCCCAACCTT
TTGTAATGCCATTTTCTAACTTATTTCTGAGTGTAAGTCTCAGCTTAAAGTTGTGTAATACT
AAAATCACGAGAACACCTAAACAACAACCAAAAATCTATTGTGGTATGCACTTGATTAACCT
ATAAAATGTTAGAGGAACTTTCACATGAATAATTTTTGTCAAATTTTATCATGGTATAATT
TGTA AAAAATAAAAAGAAATTACAAAAGAAATTATGGATTTGTCAATGTAAGTATTTGTCATA
TCTGAGGTCCAAAACCACAATGAAAGTGCTCTGAAGATTTAATGTGTTTATTCAAATGTGGT
CTCTTCTGTGTCAAATGTTAAATGAAATATAAACATTTTTTAGTTTTTAAATATTCCGTGG
TCAAATTTCTTCCTCACTATAATTGGTATTTACTTTTACCAAAAATTCTGTGAACATGTAAT
GTAAC TGGCTTTTGAGGGTCTCCCAAGGGGTGAGTGGACGTGTTGGAAGAGAGAAGCACCAT
GGTCCAGCCACCAGGCTCCCTGTGTCCCTTCATGGGAAGGTCTTCCGCTGTGCCTCTCATT
CCAAGGGCAGGAAGATGTGACTCAGCCATGACACGTGGTCTGGTGGGATGCACAGTCACTC
CACATCCACCACTG

57/330

FIGURE 57

MSGFLEGLRCSECIDWGEKRNTIASIAAGVLFFTGWIIIDAAVIYPTMKDFNHSYHACGVI
ATIAFLMINAVSNGQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILFGGYVAKEK
DIVYPGIAVFFQNAFIFFGGLVFKFGRTEDLWQ

58/330

FIGURE 58

TTCTTGGCTAAAATCGGGGGAGTGAGGCGGGCCGGCGCGGCGCGACACCGGGCTCCGGAACC
ACTGCACGACGGGGCTGGACTGACCTGAAAAAATGTCTGGATTTCTAGAGGGCTTGAGATG
CTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTAC
TATTTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGAT
TTCAACCACTCATACCATGCCTGTGGTGTATAGCAACCATAGCCTTCCTAATGATTAATGC
AGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTG
CTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGG
ATTCTTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGCTGTATT
TTTCCAGAATGCCTTCATCTTTTTTTGGAGGGCTGGTTTTTTAAGTTTGGC

59/330

FIGURE 59

TGGACGGACCTGAAAAAATGTTTGGATTTNTAGAGGGNTTGAGATGTTTCAGAATGCATGAC
TGGGGGAAAAGCGCAAATACTATTGCTTCCATTGCTGCTGGTGTANTATTTTTTACAGGCTG
GTGGATTATCATAGATGCAGNTGTTATTTATCCCACCATGAAAGATTTCAACCANTCATACC
ATGCCTGTGGTGTTATAGCAACCATAGCCTTCNTAATGATTAATGCAGTATCGAATGGACAA
GTCCGAGGTGATAGTTACAGTGAAGGTTGTTTGGGTCAAACAGGTGCTCGCATTTGGCTTTT
CGTTGGTTTCATGTTGGCCTTTGGATCTCTGATTGCATCTATGTGGATTCTTTTTTGGAGGTT
ATGTTGCTAAAGAAAAAGACATAGTATACCCTGGAATTGNTGTATTTTTTCCAGAATGCCTTC
ATCTTTTTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGANTTATGGCAGTG

60/330

FIGURE 60

GGACACCGGGTTCCGGACCAATGCANGACGGGGTGGANTGACCTGAAAAAATGTTTGGATT
TTTAGAGGGCTTGAGATGNTCAGAATGCATTGACTGGGGGAAAAGCGCAATANTATTGCTTT
CCATTGCTGCTGGTGTACTATTTTTTACAGGGTGGTGGATTATCATAGATGCAGCTGTTATT
TATCCCACCATGAAAGATTTNAACCACTCATACCATGCCTGTGGTGTTATAGCAACCATAGC
CTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTT
GTTTGGGTCAAACAGGTGNTCGCATTTGGCTTTTCGTTGGTTTCATGTTGGCCTTTGGATTT
CTGATTGNATTCTATGCGGATTCTTCTTGGAGGTTATGTTGCTAAAGAAAAGACATAGTAT
ACCCTGGAATTNCTNTATTTTTTCCAGAATGCC

61/330

FIGURE 61

TAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATANTATTGCTTCC
ATTGNTGNTGGTGTANTATTTTTTTTACAGGCTGGTGGATTATNATAGATGCAGCTGTTATTT
ATCCCACCATGAAAGATTTNAACCANTCATACCATGCCTGTGGTGTTATAGCAACCATAGCC
TTCCTAATGATTAATGCAGTATNGAATGGACAAGTCCGAGGTGATAGTTACAGTGAAGGTTG
TTTGGGTCAAACAGGTGNTNGCATTTGGCTTTTNGTTGGTTTCATGTTGGCCTTTGGATCTN
TGATTGCATTTATGTGGATTNTTTTTTGGAGGTTATGTTGCTAAAGNAAAAGACATAGTATAC
CCTGT

62/330

FIGURE 62

GGGAGGCTGTGNCCGTTTTGTTTTNTTGGCTAAAATCGGGGGAGTGAGGCGGCCCCGGCGCGG
CGNGACACCGGGTTCCGGGAACCATTCACGACGGGGTGGACTGACCTGAAAAAATGTTTG
GATTTNTAGAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATT
GCTTCCATTGCTGCTGGTGTACTATTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGT
TATTTATCCCACCATGAAAGATTTCAACCACTCATACCATGCCTGTGGTGTTATAGCAACCA
TAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGTGATAGTTACAGTGAA
GGTTGTCTGGGTCAAACAGGTGCTCGCATTGCGCTTTTCGTTGGTTTCATGTTGGCCTTTGG
ATNTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAG
TATACCCTGGAATTGCTGTATTTTTCCAGAATGCCTTCATNTTTTTTGGAGGGCTG

63/330

FIGURE 63

CGACGCCGGCGTG**ATG**TGGCTTCCGCTGGTGTGCTGCTCCTGGCTGTGCTGCTGCTGGCCGTCC
TCTGCAAAGTTTACTTGGGACTATTCTCTGGCAGCTCCCCGAATCCTTTCTCCGAAGATGTC
AAACGGCCCCCAGCGCCCCCTGGTAACTGACAAGGAGGCCAGGAAGAAGGTTCTCAAACAAGC
TTTTTCAGCCAACCAAGTGCCGGAGAAGCTGGATGTGGTGGTAATTGGCAGTGGCTTTGGGG
GCCTGGCTGCAGCTGCAATTCTAGCTAAAGCTGGCAAGCGAGTCCTGGTGTGGAACAACAT
ACCAAGGCAGGGGGGCTGCTGTCATACCTTTGGAAAGAATGGCCTTGAATTTGACACAGGAAT
CCATTACATTGGGCGTATGGAAGAGGGCAGCATTGGCCGTTTTATCTTGGACCAGATCACTG
AAGGGCAGCTGGACTGGGCTCCCCCTGTCTCTCTCTTTTGGACATCATGGTACTGGAAGGGCCC
AATGGCCGAAAGGAGTACCCCATGTACAGTGGAGAGAAAAGCCTACATTACAGGGCCTCAAGGA
GAAGTTTCCACAGGAGGAAGCTATCATTGACAAGTATATAAAGCTGGTTAAGGTGGTATCCA
GTGGAGCCCCCTCATGCCATCCTGTTGAAATTCCTCCCATTGCCCGTGGTTTACAGCTCCTCGAC
AGGTGTGGGCTGCTGACTCGTTTCTCTCCATTCTTCAAGCATCCACCCAGAGCCTGGCTGA
GGTCTCTGCAGCAGCTGGGGGCTCCTCTGAGCTCCAGGCAGTACTCAGCTACATCTTCCCCA
CTTACGGTGTCAACCCCAACACAGTGCCTTTTTCCATTGCACGCCCTGCTGGTCAACCACTAC
ATGAAAGGAGGCTTTTTATCCCCGAGGGGGTTCAGTGAAATTGCCTTCCACACCATCCCTGT
GATTACAGCGGGCTGGGGGCGCTGTCTCACAAAGGCCACTGTGCAGAGTGTGTTGCTGGACT
CAGCTGGGAAAGCCTGTGGTGTGAGTGTGAAGAAGGGGCATGAGCTGGTGAACATCTATTGC
CCCATCGTGGTCTCCAACGCAGGACTGTTCAACACCTATGAACACCTACTGCCGGGGAACGC
CCGCTGCCTGCCAGGTGTGAAGCAGCAACTGGGGACGGTGC GGCCCGGCTTAGGCATGACCT
CTGTTTTTCTGCTGCGAGGCACCAAGGAAGACCTGCATCTGCCGTCCACCAACTACTAT
GTTTACTATGACACGGACATGGACCAGGCGATGGAGCGCTACGTCTCCATGCCCAGGGAAGA
GGCTGCGGAACACATCCCCCTTCTCTTCTTCTCGCTTTCCCATCAGCCAAAGATCCGACCTGGG
AGGACCGATTCCCAGGCCGGTCCACCATGATCATGCTCATACCCACTGCCTACGAGTGGTTT
GAGGAGTGGCAGGCGGAGCTGAAGGGGAAAGCAGGAGTACTATGAGACCTTCAAAAATC
CTTTGTGGAAGCCTCTATGTGAGTGGTCTGAAACTGTTCCACAGCTGGAGGGGAAGGTGG
AGAGTGTGACTGCAGGATCCCCACTCACCACCAAGTCTATCTGGCTGCTCCCCGAGGTGCC
TGCTACGGGGCTGACCATGACCTGGGCGCCTGCACCCCTTGTGTGATGGCCTCCTTGAGGGC
CCAGAGCCCCATCCCCAACCTCTATCTGACAGGCCAGGATATCTTCACTGTGCTAGTGGTGG
GGCCCTGCAAGGTGCCCTGCTGTGCAGCAGCGCCATCCTGAAGCGGAACCTGTACTCAGAC
CTTAAGAATCTTGATTCTAGGATCCGGGACAGAAAGAAAAGAAT**TAG**TTCCATCAGGGAGG
AGTCAGAGGAATTTGCCCAATGGCTGGGGCATCTCCCTTGACTTACCCATAATGTCTTTCTG
CATTAGTTCTTGCACGTATAAAGCACTCTAATTTGGTTCTGATGCCTGAAGAGAGGCCTAG
TTTAAATCACAAATCCGAATCTGGGGCAATGGAATCACTGCTTCCAGCTGGGGCAGGTGAGA
TCTTTACGCCTTTTATAACATGCCATCCCTACTAATAGGATATTGACTTGGATAGCTTGATG
TCTCATGACGAGCGGCGCTCTGCATCCCTCACCCATGCCCTCCTAACTCAGTGATCAAAGCGA
ATATTCCATCTGTGGATAGAACCCCTGGCAGTGTGTGCTCAGCTCAACCTGGTGGGTTTCAAGG
TGTCCTGAGGCTTCTGCTCTCATTCATTTAGTGCTACGCTGCACAGTTCTACACTGTCAAGG
GAAAAGGGAGACTAATGAGGCTTAACTCAAACCTGGGCGTGGTTTTTGGTTGCCATTCCATA
GGTTTGGAGAGCTCTAGATCTCTTTTGTGCTGGGTTTCACTGGCTCTTCAGGGGACAGGAAAT
GCCTGTGTCTGGCCAGTGTGGTTCTGGAGCTTTGGGGTAACAGCAGGATCCATCAGTTAGTA
GGGTGCATGTGAGATGATCATATCCAATTCATATGGAAGTCCCGGGTCTGTCTTCTCTTATCA
TCGGGGTGGCAGCTGGTTCTCAATGTGCCAGCAGGGACTCAGTACCTGAGCCTCAATCAAGC
CTTATCCACCAAATACACAGGGAAGGGTGTGTCAGGGAAGGGTGACATCAGGAGTCAGGGCA
TGGACTGGTAAGATGAATACTTTGCTGGGCTGAAGCAGGCTGCAGGGCATTCCAGCCAAGGG
CACAGCAGGGGACAGTGCAGGGAGGTGTGGGGTAAGGGAGGGAAGTCACATCAGAAAAGGGA
AAGCCACGGAATGTGTGTGAAGCCCAGAAATGGCATTTGCAGTTAATTAGCACATGTGAGGG
TTAGACAGGTAGGTGAATGCAAGCTCAAGGTTTGGAAAAATGACTTTTTCAGTTATGTCTTTG
GTATCAGACATACGAAAGGTCTCTTTGTAGTTTCGTGTTAATGTAACATTAATAAATTTATTG
ATTCCATTGCTTTAAAAA

64/330

FIGURE 64

MWLPLVLLLAVLLLAVLCKVYLGLFSGSSPNPFSEDVKRPPAPLVTDKEARKKVLKQAFSAN
QVPEKLDVVVIGSGFGGLAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLEFDTGIHYIG
RMEEGSIGRFILDQITEGQLDWAPLSSPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQ
EEAIIIDKYIKLVKVSSGAPHAILLKFLPLPVVQLLDRCGLLTRFSPFLQASTQSLAEVLQQ
LGASSELQAVLSYIFPTYGVTPNHSAFSMHALLVNHYMKGGFYPRGGSSEIAFHTIPVIQRA
GGAVLTKATVQSVLLDSAGKACGVSVKKGHELVNIYCPIVVSNAGLFNTYEHLLPGNARCLP
GVKQQLGTVRPGLGMTSVFICLRGTKEDLHLPSTNYYVYYDTDMDQAMERYVSMPREEAAEH
IPLLFFAFPSAKDPTWEDRFPGRSTMIMLIPTAYEWFEWQAELKGKRGSDYETFKNSFVEA
SMSVVLKLFQLEGKVESVTAGSPLTNQFYLAAPRGACYGADHDLGRLHPCVMASLRAQSPI
PNLYLTGQDIFTCGLVGALQGALLCSSAILKRNLYSDLKNLDSRIRAQKKKN

65/330

FIGURE 65

[illegible]

66/330

FIGURE 66

MRVRIGLTLLLCVLLSLASASSDEEGSQDES LDSKTTLTSDSVKDHTTAGRVVAGQIFLD
SESELESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEP KVKVRKPALTAIEGTAHG
EPCHF PFLFLDKEYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEEAAKRRQM QEAEMM
YQTGMKILNGSNKKSQKREAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMF EK
LTEEGSPKGQTALGFLYASGLGVNSSQAKALVYYTFGALGGNLI AHMVLVSRL

67/330

FIGURE 67

CTTCCCAGCCCTGTGCCCCAAAGCACCTGGAGCATATAGCCTTGCAGAACTTCTACTTGCCT
GCCTCCCTGCCTCTGGCC**ATG**GCCTGCCGGTGCCTCAGCTTCCTTCTGATGGGGACCTTCCT
GTCAGTTTCCCAGACAGTCCTGGCCCAGCTGGATGCACTGCTGGTCTTCCCAGGCCAAGTGG
CTCAACTCTCCTGCACGCTCAGCCCCCAGCACGTCACCATCAGGGACTACGGTGTGTCCTGG
TACCAGCAGCGGGCAGGCAGTGGCCCTCGATATCTCCTCTACTACCGCTCGGAGGAGGATCA
CCACCGGCCTGCTGACATCCCCGATCGATTCTCGGCAGCCAAGGATGAGGCCCAATGCCT
GTGTCCTCACCATTAGTCCCGTGCAGCCTGAAGACGACGCGGATTACTACTGCTCTGTTGGC
TACGGCTTTAGTCCC**TAG**GGGTGGGGTGTGAGATGGGTGCCTCCCCTCTGCCTCCCATTTCT
GCCCCTGACCTTGGGTCCCTTTTAAACTTTCTCTGAGCCTTGCTTCCCCTCTGTAAAATGGG
TTAATAATATTCAACATGTCAACAAC

68/330

FIGURE 68

MACRCLSFLMGTFLSVSQTVLAQLDALLVFPQVAQLSCTLSPQHVTIRDYGVSWYQQRAG
SAPRYLLYYRSEEDHHRPADIPDRFSAKDEAHNACVLTISPVPEDDADYYCSVGYGFSF

69/330

FIGURE 69

GCCGCCCCGCCCCGAGACCGGGCCCCGGGGGCGCGGGGCGGCGGGGATGCGGCGCCCCGGGGCGG
CGATGACCGCGGGAGCGCACGCCGCGGGGCCCGGCCCTGACCCCGCCGCCCGCCCCGCTGAGCCC
CCCCCGGAGGTCCGGACAGGCCGAG**ATG**ACGCCGAGCCCCCTGTTGCTGCTCCTGCTGCCGC
CGCTGCTGCTGGGGGCCCTTCCCACCGGCCGCCGCCGCCGAGGCCCCCAAAGATGGCGGAC
AAGGTGGTCCCACGGCAGGTGGCCCCGGCTGGGCCGCACTGTGCGGCTGCAGTGCCCAAGTGA
GGGGGACCCGCGCGCCGCTGACCATGTGGACCAAGGATGGCCGCACCATCCACAGCGGCTGGA
GCCGCTTCCGCGTGCTGCCGCAGGGGCTGAAGGTGAAGCAGGTGGAGCGGGAGGATGCCGGC
GTGTACGTGTGCAAGGCCACCAACGGCTTCGGCAGCCTGAGCGTCAACTACACCTCGTCGT
GCTGGATGACATTAGCCCAGGGAAGGAGAGCCTGGGGCCCCGACAGCTCCTCTGGGGGTCAAG
AGGACCCCGCCAGCCAGCAGTGGGCACGACCGCGCTTCACACAGCCCTCCAAGATGAGGCGC
CGGGTGATCGCACGGCCCCGTGGGTAGCTCCGTGCGGCTCAAGTGCGTGGCCAGCGGGCACCC
TCGGCCCCGACATCACGTGGATGAAGGACGACCAAGGCCCTTGACGCGCCAGAGGCCGCTGAGC
CCAGGAAGAAGAAGTGGACACTGAGCCTGAAGAACCTGCGGCCGGAGGACAGCGGCAAATAC
ACCTGCCGCGTGTCGAACCGCGCGGCGCCATCAACGCCACCTACAAGGTGGATGTGATCCA
GCGGACCCGTTCCAAGCCCGTGCTCACAGGCACGCACCCCGTGAACACGACGGTGGACTTCG
GGGGGACCACGTCCTTCCAGTGCAAGGTGCGCAGCGACGTGAAGCCGGTGATCCAGTGGCTG
AAGCGCGTGAGTACGGCGCCGAGGGCCGCCACAACCTCCACCATCGATGTGGGCGGCCAGAA
GTTTGTGGTGCTGCCACGGGTGACGTGTGGTTCGCGGCCCGACGGCTCCTACCTCAATAAGC
TGCTCATCACCCGTGCCCGCCAGGACGATGCGGGCATGTACATCTGCCTTGGCGCCAAACAC
ATGGGCTACAGCTTCCGCAGCGCCTTCCTCACCGTGCTGCCAGACCCAAAACCGCCAGGGCC
ACCTGTGGCCTCCTCGTCCTCGGCCACTAGCCTGCCGTGGCCCGTGGTTCATCGGCATCCAG
CCGGCGCTGTCTTCATCCTGGGCACCCCTGCTCCTGTGGCTTTGCCAGGCCCAGAAGAAGCCG
TGACCCCCGCGCTGCCCTCCCCCTGCCTGGGACCCGCCCGCGGGGACGGCCCGCGACCC
CAGCGGAGACAAGGACCTTCCCTCGTTGGCCGCCCTCAGCGCTGGCCCTGGGTGTGGGCTGT
GTGAGGAGCATGGGTCTCCGGCAGCCCCCAGCACTTACTGGGCCCAGGCCAGTTGCTGGC
CCTAAGTTGTACCCCAAACCTCTACACAGACATCCACACACACACACACACACTCTCACAC
ACACTCACACGTGGAGGGCAAGGTCCACCAGCACATCCACTATCAGTGCT**TAG**ACGGCACCGT
ATCTGCAGTGGGCACGGGGGGGCCCGGCCAGACAGGCAGACTGGGAGGATGGAGGACGGAGCT
GCAGACGAAGGCAGGGGACCCATGGCGAGGAGGAATGGCCAGCACCCAGCGAGTCTGTGTG
TGAGGCATAGCCCCCTGGACACACACACACAGACACACACACTACCTGGATGCATGTATGCAC
ACACATGCGCGCACACGTGCTCCCTGAAGGCACACGTACGCACACGCACATGCACAGATATG
CCGCCTGGGCACACAGATAAGCTGCCCAAATGCACGCACACGCACAGAGACATGCCAGAACA
TACAAGGACATGCTGCCTGAACATACACGCACACCCATGCGCAGATGTGCTGCCTGGACA
CACACACACACACGAGATATGCTGTCTGGACGCACACACGTGCAGATATGGTATCCGGACACA
CACGTGCACAGATATGCTGCCTGGACACACAGATAATGCTGCCTTGACACACACATGCACGG
ATATTGCCTGGACACACACACACACACACGCGTGACAGATATGCTGTCTGGACACGCACAC
ACATGCAGATATGCTGCCTGGACACACACTTCCAGACACACGTGCACAGGCGCAGATATGCT
GCCTGGACACACGCAGATATGCTGTCTAGTCACACACACACGCAGACATGCTGTCCGGACAC
ACACACGCATGCACAGATATGCTGTCCGGACACACACACGCACGCAGATATGCTGCCTGGAC
ACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGATATTGCCTGGACACACACA
TGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATATGCTGTCCGGATACACACG
CACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGACACACATGCACACACAGGT
GCAGATATGCTGCCTGGACACACACAGATAATGCTGCCTCAACACTCACACACGTGCAGATA
TATTGCCTGGACACACACATGTGCACAGATATGCTGTCTGGACATGCACACACGTGCAGATA
TGCTGTCCGGATACACACGCACGCACACATGCAGATATGCTGCCTGGGCACACACTTCCGGA
CACACATGCACACACAGGTGCAGATATGCTGCCTGGACACACGCAGACTGACGTGCTTTTGG
GAGGTGTGCCGTGAAGCCTGCAGTACGTGTGCCGTGAGGCTCATAGTTGATGAGGGACTTT
CCTGCTCCACCGTCACTCCCCCAACTCTGCCCGCCTCTGTCCCCGCCTCAGTCCCCGCCTC
CATCCCCCCTGTGTCCCCCTGGCCTTGCCGCTATTTTGGCACCTGCCTTGGGTGCCCAGG
AGTCCCCTACTGCTGTGGGCTGGGGTTGGGGGCACAGCAGCCCCAAGCCTGAGAGGCTGGAG
CCCATGGCTAGTGGCTCATCCCCAGTGCATTCTCCCCCTGACACAGAGAAGGGGCCTTGTA
TTTATATTTAAGAAATGAAGATAATATTAATAATGATGGAAGGAAGACTGGGTTGCAGGGAC
TGTGGTCTCTCCTGGGGCCCCGGGACCCGCTGGTCTTTTTCAGCCATGCTGATGACCAACCCCC
GTCCAGGCCAGACACACCCCCCAGCTGTGCTGGTGGGCCAGATCTGTGTAATTTTA
TGTAGAGTTTGAGCTGAAGCCCCGTATATTTAATTTATTTTGTAAACACAAAA

70/330

FIGURE 70

MTPSPLLLLLLLPPLLLGAFPPAAAARGPPKMADKVVPRQVARLGRTVRLQCPVEGDPPPLTM
WTKDGRTIHSWWSRFRVLPQGLKVKQVEREDAGVYVCKATNGFGSLSVNYTLVVLDISP
ESLGPDSSSSGGQEDPASQQWARPRFTQPSKMRRRVIA RPVGSSVRLKCVASGHPRPDITWMK
DDQALTRPEAAEPRKKKWTLSLKNLRPEDSGKYTCRVSNRAGAINATYKVDVIQRTRSKPVL
TGTHPVNTTVDFGGTTSFQCKVRSDVKPVIQWLKRVEYGAEGRHNSTIDVGGQKFVVLPTGD
VWSRPDGSYLNKLLITRARQDDAGMYICLGANTMGYSFRSAFLTVLPDPKPPGPPVASSSSA
TSLPWPVVIGIPAGAVFILGTLLLWLCQAQKKPCTPAPAPPLPGHRPPGTARDRSGDKDLPS
LAALSAGPGVGLCEEHGSPAAPQHLLGPGPVAGPKLYPKLYTDIHTHTHTSHHTSHVEGKV
HQHIHYQC

71/330

FIGURE 71

CCCAGCTGAGGAGCCCTGCTCAAGACACGGTCACTGGATCTGAGAACTTCCCAGGGGACCGCATTCCAGAGTC
AGTGACTCTGTGAAGCACCCACATCTACCTCTTGCCACGTTCCACAGGGCTTGGGGGAAAG**ATGGT**GGGGACCA
AGGCTGGGTGTTCTCCTTCTGGTCTGGAAGTCACATCTGTGTTGGGGAGACAGATGCTCACCCAGTCA
GTAAGAAGAGTCCAGCCTGGGAAGAAGAACCACAGCATCTTTGCCAAGCCTGCCGACACCCTGGAGAGCCCTGG
TGAGTGGACAACATGGTTCAACATCGACTACCCAGGCGGGAAGGCGACTATGAGCGGCTGGACGCCATTTCGCT
TCTACTATGGGGACCGTGTATGTGCCCGTCCCTGCGGCTAGAGGCTCGGACCACTGACTGGACACCTGCGGGC
AGCACTGGCCAGGTGGTCCATGGTAGTCCCCGTGAGGGTTTCTGGTGCCTCAACAGGGAGCAGCGGCCTGGCCA
GAACTGCTCTAATTACACCGTAGCCTTCTGCTGCCACCAAGATCCCTGCGCCGAGACACAGAGCGGCATCTGGA
GCCCATGGTCTCCCTGGAGCAAGTGTCTGAGTGCCTGTGGTCAAGTGGGTCAGACTGGGGTCCAGACTCGCACACGCATTTGC
TTGGCAGAGATGGTGTGCTGTGAGTGAAGGACAGCAAGAGGGTCAAGCACTGCATGGGCCAGGACTGTACAGC
CTGTGACCTGACCTGCCAATGGGCGAGTGAATGCTGACTGTGATGCCTGCATGTGCCAGGACTTCATGCCTT
ATGGGGCTGTCTCCCTTCCCGGAGGTGCCCCAGCCTCAGGGGCTGCTATCTACCTCCTGACCAAGACGCCGAAG
CTGCTGACCCAGACAGACAGTGTGGGAGATTCCGAATCCCTGGCTTGTGCCCTGATGGCAAAGCATCTTGAA
GATCACAAAGGTCAAGTTTGCCCCATTGTACTACAATGCCAAGACTAGCCTGAAGGACACCACTCAAGG
CAGAGTTTGTGAGGGCAGAGACTCCATACATGGTGTGAACCTGAGACAAAAGCAGGAGAGCTGGGCAGAGC
GTGTCTGTGCTGTGAAGGCCACAGGGAAGCCAGGCCAGACAAAGTATTTTGGTATCATAATGACACATTGCA
GGATCCTTCCCTTACAAAGCATGAGAGCAAGTGGTGTGAGGAACTGCAGCAGCAGGATAATGGGATGACTGGCT
TTTGCAAGGCCCAGAGTGTGCTGGGGCTGTGAAGTCCAAGGTTGCCAGCTGATTGTACAGCATCTGATGAG
ACTCCTTGCAACCCAGTTCTGTGAGAGCTATCTTATCCGGCTGCCCATGATTGCTTTGAGAATGCCACCACTC
CTGTACTACTGACCTGCCAATGGGCGAGTGCCTTGTAAAGCTGTGACAGGAGCAGGATAATGGGATGCTGCTG
ATGCTGTGCAGAACTGCTGTGGCATCTCCAAGACAGAGGAAAGGGAGATCCAGTGCAGTGGCTACACGCTACCC
ACCAAGGTGGCCAAGGAGTGCAGTGCAGCGGTGTACGGAACCTCGGAGCATCGTGGGGGCGCTGTGCTGAGC
TGCTGACAATGGGGAGCCCATGCGCTTGGCCATGTGTACATGGGGAACAGCCGTGTAAGCATGACTGGCTACA
AGGGCACTTTTACCCTCCATGTCCCCCAGGACACTGAGAGGCTGGTGTACATTTGTGGACAGGCTGCAGAAG
TTTGTCAACACCACCAAGTGTACCTTTCAACAAGAAGGGGAGTGCCGTGTTCCATGAAATCAAGATGCTTCG
TCGGAAGAGAGCCATCTTTGGAAGCCATGGAGCAACATCATCCCCCTGGGGGAAGTGGTTGGTGAAGACC
CCATGGCTGAAGTGGAGATTCCATCCAGGAGTTTCTACAGGCAGAATGGGGAGCCCTACATAGGAAAAGTGAAG
GCCAGTGTGACCTTCTGGATCCCCGGGAATATTCCACAGCTCAGCTGCCAGACTGACCTGAACCTCATCAA
TGACGAAGGAGACACTTTCCCCCTCGGACGTATGGCATGTTCTCTGTGGACTTCAAGATGAGGTACCTCAG
AGCCACTTAATGCTGGCAAAGTGAAGGTCCACCTTGACTCGACCCAGGTCAAGATGCCAGAGCACATATCCACA
GTGAAACTCTGGTCACTCAATCCAGACACAGGGCTGTGGGAGGAGGAAGGTGATTTCAAATTTGAAAATCAAAG
GAGGAACAAAAGAGACAGCTTTTCCAGATTAGCATGGCCAAGCCAAGGCCAACTCAGCTGAGGAGAGCAATGGGCCA
TCTATGCCCTTTGAGAACCTCCGGGCATGTGAAGAGGCACCAACCCAGTGCAGCCCACTTCCGGTTCTACCAGATT
GAGGGGGATCGATATGACTACAACACAGTCCCCCTCAACGAAGATGACCTATGAGCTGGACTGAAGACTATCT
GGCATGGTGGCCAAAGCCGATGGAATTGAGGGCTGCTATATCAAGGTGAAGATTGTGGGGCACTGGAAGTGA
ATGTGCGATCCCGCAACATGGGGGGCACTCATCGGCGGACAGTGGGGAAGCTGTATGGAATCCGAGATGTGAGG
AGCACTCGGGACAGGGACAGCCCAATGTCTCAGTGCCTGTCTGGAGTTCAAGTGCAGTGGGATGCTCTATGA
TCAGGACCGTGTGGACCGCACCTGGTGAAGGTCAATCCCCCAGGGCAGCTGCCGTGAGCCAGTGTGAACCCCA
TGCTGCATGAGTACCTGGTCAACCACTTGCCACTTGCACTCAACAACGACACCAAGTGAATACCATGCTGGCA
CCCTTGGACCCACTGGGCCACAACATATGGCATCTACACTGTCACTGACCAGGACCTCGCACGGCCAAGGAGAT
CGCGCTCGGCGGTGCTTTGATGGCACATCCGATGGCTCCTCCAGAATCATGAAGAGCAATGTGGGAGTAGCCC
TCACCTTCAACTGTGTAGAGAGGCAAGTAGGCCGCCAGAGTGCCTTCCAGTACCTCAAAGCACCCAGCCAG
TCCCTTGCTGCAGGCACTGTCCAAGGAAGAGTGCCCTCGAGGAGGCAGCAGCGAGCGAGCAGGGGTGGCCAGCG
CCAGGGTGGAGTGGTGGCTCTCTGAGATTTCTAGAGTTGCTCAACAGCCCTGATCAAC**TAA**GTTTTGTGGT
ACTTACCCTCTTCTGCCCTCATTTTCTGTGACAGCCATTGTGAGACTGATGCACAACTGTCACTTGGTTAAT
TTAAGCACTTCTGTTTTCTGTGAATTTGCTTGTGTTTCTTCTATGCCTTTACTTACTTTGTCCCATGCTACTGA
TTGGCACGTGGCCCCACAATGGCACAAATAAAGCCCTTTGTGAAACTGTTCTTTAAATGAAACACAAGAAAT
GGCCACTGGTAAAACTCTGCAGCTTCAACTGTACTTCAATTAATGCCATTAATGCAATATACTTCTCTTCTT
TTTGCATGGTTTTTGGCCACCTCTGCAATAGTGATAATCTGATGCTGAAGATCAAATAACCAATATAAAGCATAT
TTCTTGGCCTTGTCTCCACAGGACATAGGCAAGCCTTGTATCATAGTTTCATACATATAAATGGTGGTGAATAAAG
AAATAAAACACAATACTTTTACTTGAAATGTAAATAACTTATTTATTTCTTTGCTAAATTTGGAATTTCTAGTGC
ACATTCAAAGTTAAGCTATTAATATAGGGTGTATCATAGTTCCCTCTACCAAGTCTGGAAAGAACATCTCCTGGT
ATCCACAATTACACAGGTTGCTAAGTGTATTTGTACATTTCCCTTTGCTTTGCTTTGCTTAGAAAC
CCAGTGTAGCCAGGGCAGATGTCAATAAATGCATACTCTGTATTTTCAAAAAA

72/330

FIGURE 72

MVGTKAWVFSFLVLEVTSVLGRQTMLTQSVRRVQPGKKNPSIFAKPADTLESPGEWTTWFNI
DYPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQ
RPGQNCSNYTVRFLCPPGSLRRDTERIWSPPWSPWSKCSAACGQTGVQTRTRICLAEMVSLCS
EASEEGQHCMGQDCTACDLTCPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAYLLTK
TPKLLTQTDSDGRFRI PGLCPDGKSILKITKVKFAPIVLTMPKTSLSKAATIKAEFVRAETPY
MVMNPETKARRAGQSVSLCCKATGKPRPDKYFWYHNDTLLDPSLYKHESKLVLRKLQHQAG
EYFCKAQSDAGAVKSKVAQLIVTASDETPCNPVPESYLIRLPHDCFQONATNSFYDVGRCPV
KTCAGQQDNGIRCRDAVQNCCGISKTEEREIQCSGYTLPTKVAKECSCQRCTETRSIVRGRV
SAADNGEPMRFHGVYMGNSRVSMGTGYKGTFTLHVPQDTERLVLT FVDRLQKFVNTTKVLPFN
KKGS AVFHEIKMLRRKEPITLEAMETNIIPLGEVVGEDPMAELEIPSRSFYRQNGEPYIGKV
KASVTFLDPRNISTATAAQTDLNFINDGDTFPLRITYGMFSVDFRDEVTSEPLNAGKVKVHL
DSTQVKMPEHISTVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLFNL
DVPESRRRCFVKVRAYRSEFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGA
CVPAFCDDQSPDAYSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLNRYRRTDHEDPR
VKKTAFQISMAKPRPNNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYNTPVFN
EDDPMSTEDYLAWWPKPMEFRACYIKVKIVGPLEVNVRSRNMGGTHRRTVGKLYGIRDVRS
TRDRDQPNVSAACLEFKCSGMLYDQDRVDRTL VKVIPQGSCRRASVNPMLHEYLVNHLPLAV
NNDTSEYTMLAPLDPLGHNYGIYTVTDQDPRTAKEIALGRCFDGTS DGSSRIMKSNVGVALT
FNCVERQVGRQSAFQYLQSTPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFPRVA
QQPLIN

73/330

FIGURE 73

CTGCAAGTTGTTAACGCCTAACACACAAGTATGTTAGGCTTCCACCAAAGTCCTCAATATACCTGAATACGCAC
AATATCTTAACTCTTCATATTTGGTTTTGGGATCTGCTTTGAGGTCCCATCTTCATTTAAAAAAAATACAGAG
ACCTACCTACCCGTACGCATACATACATATGTGTATATATATGTAAACTAGACAAAGATCGCAGATCATAAAGC
AAGCTCTGCTTTAGTTTTCCAAGAAGATTACAAAGAATTTAGAG**ATGT**ATTTGTCAAGATCCCTGTCGATTCATG
CCCTTTGGGTTACGGTGTCTCAGTGATGCAGCCCTACCCTTTGGTTTTGGGGACATTATGATTTGTGTAAGACT
CAGATTTACACGGAAGAAGGGAAAGTTTGGGATTACATGGCCTGCCAGCCGGAATCCACGGACATGACAAAATA
TCTGAAAGTGAAACTCGATCCTCCGGATATTACCTGTGGAGACCCTCCTGAGACGTTCTGTGCAATGGGCAATC
CCTACATGTGCAATAATGAGTGTGATGCGAGTACCCCTGAGCTGGCACACCCCCCTGAGCTGATGTTTGATTTT
GAAGGAAGACATCCCTCCACATTTTGGCAGTCTGCCACTTGGAAGGAGTATCCCAAGCCTCTCCAGGTTAACAT
CACTCTGTCTTGGAGCAAAACCATTGAGCTAACAGACAACATAGTTATTACCTTTGAATCTGGGCGTCCAGACC
AAATGATCCTGGAGAAGTCTCTCGATTATGGACGAACATGGCAGCCCTATCAGTATTATGCCACAGACTGCTTA
GATGCTTTTTACATGGATCCTAAATCCGTGAAGGATTTATCACAGCATACGGTCTTAGAAATCATTTGCACAGA
AGAGTACTCAACAGGGTATACAACAAATAGCAAATAATCCACTTTGAAATCAAAGACAGGTTTCGCGCTTTTTG
CTGGACCTCGCCTACGCAATATGGCTTCCCTCTACGGACAGCTGGATACAACCAAGAACTCAGAGATTTCTTT
ACAGTCACAGACCTGAGGATAAGGCTGTTAAGACCAGCCGTTGGGGAAATATTTGTAGATGAGCTACACTTGGC
ACGCTACTTTTTACGCGATCTCAGACATAAAGGTGCGAGGAAGGTGCAAGTGTAAATCTCCATGCCACTGTATGTG
TGTATGACAACAGCAAATTGACATGCGAATGTGAGCACAACTACAGGTCCAGACTGTGGGAAATGCAAGAAG
AATTATCAGGGCCGACCTTGGAGTCCAGGCTCCTATCTCCCCATCCCCAAAGGCACTGCAAATACCTGTATCCC
CAGTATTTCCAGTATTGGTACGAATGTCTGCGACAACGAGCTCCTGCACTGCCAGAACGGAGGGACGTGCCACA
ACAACGTGCGCTGCCTGTGCCCCGGCCGCATACACGGGCATCCTCTGCGAGAAGCTGCGGTGCGAGGAGGCTGGC
AGCTGCGGCTCCGACTCTGGCCAGGGCGCGCCCCGACGGCACCCAGCGCTGCTGCTGCTGACCACGCTGCT
GGGAACCGCCAGCCCCCTGGTGTCT**TAGGT**GTACCTCCAGCCACACCGGACGGGCTGTGCCGTGGGGAAGCA
GACACAACCCAAACATTTGCTACTAACATAGGAAACACACACATACAGACACCCCCACTCAGACAGTGTACAAA
CTAAGAAGGCCTAACTGAACTAAGCCATATTTATCACCCGTGGACAGCACATCCGAGTCAAGACTGTTAATTTT
TGACTCCAGAGGAGTTGGCAGCTGTTGATATTATCACTGCAAATCACATTGCCAGCTGCAGAGCATATTGTGGA
TTGGAAAGGCTGCGACAGCCCCCAAACAGGAAAGACAAAAACAAACAAATCAACCGACCTAAAAACATTGGC
TACTCTAGCGTGGTGCGCCCTAGTACGACTCCGCCAGTGTGTGGACCAACCAAATAGCATTCTTTGCTGTCTAG
GTGCATTGTGGGCATAAGGAAATCTGTTACAAGCTGCCATATTGGCCTGCTTCCGTCCCTGAATCCCTTCCAAC
CTGTGCTTTAGTGAACGTTGCTCTGTAACCCCTCGTTGGTTGAAAGATTTCTTTGCTGATGTTAGTGATGCACA
TGTGTAACAGCCCCCTCTAAAAGCGCAAGCCAGTCATACCCCTGTATATCTTAGCAGCACTGAGTCCAGTGCGA
GCACACACCCACTATACAAGAGTGGCTATAGGAAAAAGAAAGTGTATCTATCCTTTTGTATTCAAATGAAGTT
ATTTTCTTGAACACTGTAATATGTAGATTTTTTGTATTATTGCCAATTTGTGTTACCAGACAATCTGTTAAT
GTATCTAATTCGAATCAGCAAAGACTGACATTTTATTTTGTCTCTTTCTGTTCTGTTTGTTCCTGCTGTCAGA
GATTTCTCTGTAAGGGCAACGAACGTGCTGGCATCAAAGAATATCAGTTTACATATATAACAAGTGAATAAGA
TTCCACCAAAGGACATTCTAAATGTTTTCTTGTGCTTTAACACTGGAAGATTTAAAGAATAAAAACTCCTGCA
TAAACGATTTTCAAGAAATTTGTATTGCAATTTCTTAAGATGAAAGGAACAGCCACCAAGCAGTTTACACTCACT
TTACTGATTTCTGTGTGGACTGAGTACATTCAGCTGACGAATTTAGTTCCAGGAAGATGGATTGATGTTCACT
AGCTTGGACAACCTCTGCAAAATATGAGACTATTTCCACTTGGGAAAAATTACAACAGCAAAAAAAAAAAAAA
AAAAAA

74/330

FIGURE 74

MYLSRSLSIHALWVTVSSVMQPYPLVWGHYDLCKTQIYTEEGKVWDYMACQPESTDMTKYLK
VKLDPPDITCGDPPETFCAMGNPYMCNNECDASTPELAHPPELMFDFEGRHPSTFWQSATWK
EYPKPLQVNITLSWSKTIELTDNIVITFESGRPDQMILEKSLDYGRTWQPYQYYATDCLDAF
HMDPKSVKDLSQHTVLEIICTEEYSTGYTTNSKIIHFEIKDRFALFAGPRLRNMASLYGQLD
TTKKLRDFFTVDLRLRLLRPAVGEIFVDELHLARYFYAISDIKVRGRCKCNLHATVCVYDN
SKLTCECEHNTTGPDCGKCKKNYQGRPWSPGSYLPIPKGTANTCIPSISSIGTNVCDNELLH
CQNGGTCHNNVRCLCPAAYTGILCEKLRCEEAGSCGSDSGQGAPPHGTPALLLLTTLLGTAS
PLVF

75/330

FIGURE 75

CCCACGCGTCCGGGTGACCTGGGCCGAGCCCTCCCGGTCGGCTAAGATTGCTGAGGAGGCGG
CGGGTAGCTGGCAGGCGCCGACTTCCGAAGGCCGCGTCCGGGCGAGGTGTCCTCATGACTT
CTCTTGTGGACC**ATG**TCCGTGATCTTTTTTGCCTGCGTGGTACGGGTAAGGGATGGACTGCC
CCTCTCAGCCTCTACTGATTTTTTACCACACCCAAGATTTTTTGAATGGAGGAGACGGCTCA
AGAGTTTAGCCTTGC GACTGGCCCAGTATCCAGGTCGAGGTTCTGCAGAAGGTTGTGACTTT
AGTATACATTTTTCTTCTTTCGGGGACGTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCC
AGCAGCCATGGCCTTCTGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCCTATGACA
CTACCTGCATTGGCCTAGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTAG
AAAGTGAAGTGGCATTTTAACTATGTAAGTTCCTCTCAGATGGAGTGCAGCTTGGA AAAAAT
TCAGGAGGAGCTCAAGTTGCAGCCTCCAGCGGTTCTCACTCTGGAGGACACAGATGTGGCAA
ATGGGGTGATGAATGGTCACACACCGATGCACTTGGAGCCTGCTCCTAATTTCCGAATGGAA
CCAGTGACAGCCCTGGGTATCCTCTCCCTCATTCTCAACATCATGTGTGCTGCCCTGAATCT
CATTTCGAGGAGTTCACCTTGCAGAACATTCTTTACAGGATCCAAGGAGCTGGTTCTGCTGGT
TGGACCAAACCTCG**TGA**GCCAGCCACCCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCAT
CCGGGAGCAGTGATGTCAAACCTCTGCTGCTGGGGAAATCTCATCAGCAGGGAGCCTGTGGA
AAAGGGCATGTCAGTGAAATCTGGGAATGGCTGGATTTCGGAAACATCTGCCCATGTGTATTG
ATGGCAGAGCTGTTGCCCACAAGCGCCTTTTATTTAGGGTAAAATTAACAAATCCATTCTAT
TCCTCTGACCCATGCTTAGTACATATGACCTTTAACCCTTACATTTATATGATTCTGGGGTT
GCTTCAGAAGTGTTATTTTCATGAATCATTCATATGATTTGATCCCCCAGGATTCTATTTTGT
TTAATGGGCTTTTCTACTAAAAGCATAAAATACTGAGGCTGATTTAGTCAGGGCAAAACCAT
TTACTTTACATATTCGTTTTCAATACTTGCTGTTTACGTACACAAGCTTCTTACGGTTTTTC
TTGTAACAATAAATATTTTGAGTAAATAATGGGTACATTTTAACAAACTCAGTAGTACAACC
TAAACTTGATATAAAAGTGTGTAAAAATGTATAGCCATTTATATCCTATGTATAAATTAAATG
AGGTGGCTTCAGAAATGGCAGAATAAATCTAAAGTGTTTATTAAAAA AAAAAAAAAAAAAA
AAAAG

76/330

FIGURE 76

MSVIFFACVVRVRDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFS IHF
SSFGDVACMAICSCQCPAAMAFCFLET LWWEFTASYDTTCIGLASRPYAFLEFDSIIQKV KW
HFNYVSSSQMECSLEKIQEELKLQPPAVLTLED TDVANGVMNGHTPMHLEPAPNFRMEPVTA
LGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

77/330

FIGURE 77

TGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCNTATGACACTACCTGCATTGGCNT
AGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTCAGAAAGTGAAGTGGCATT
TTAACTATGTAAGTTCCTNTCAGATGGAGTGCAGCTTGAAAAAATTCAGGAGGAGCTCAAG
TTGCAGCCTCCAGCGGTTCTCANTATGGAGGACACAGATGTGGCAAATGGGGT

78/330

FIGURE 78

CTCAGCGGGCGCTTCCTCGTAGCGAGCCTAGTGGCGGGTGTTTGCATTGAAACGTGAGCGCGA
CCCGACCTTAAAGAGTGGGGAGCAAAGGGAGGACAGAGCCCTTTAAAACGAGGCGGGTGTTG
CCTGCCCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTT
TCTGTGCGCAGGCTGCGAGGAAAGGCCCTAGGCTGGGTCTGGGTGCTTGGCGGCGGCGGCTT
CCTCCCCGCTCGTCCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTA
TGGAAAGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGC
GAGTGTATTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGAC
CCGCTTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGA
TTGCGCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCTGCTCCTGCCC
TTCTCCATCATCAGCAATGAGGTGCTGCTCTCCCTGCCTCGGAACACTACTACATCCAGTGGCT
CAACGGCTCCCTCATCCATGGCCTCTGGAACCTTGTTTTTCTCTTCCCCAACCTGTCCCTCA
TCTTCCTCATGCCCTTTGCATATTTCTTCACTGAGTCTGAGGGCTTTGCTGGCTCCAGAAAG
GGTGTCTGCTGGGCGGGTCTATGAGACAGTGGTGATGTTGATGCTCCTCACTCTGCTGGTGCT
AGGTATGGTGTGGGTGGCATCAGCCATTGTGGACAAGAACAAGGCCAACAGAGAGTCACTCT
ATGACTTTTGGGAGTACTATCTCCCCACCTCTACTCATGCATCTCCTTCCTTGGGGTTCTG
CTGCTCCTGGTGTGTACTCCACTGGGTCTCGCCCGCATGTTCTCCGTCACTGGGAAGCTGCT
AGTCAAGCCCCGGCTGCTGGAAGACCTGGAGGAGCAGCTGTACTGCTCAGCCTTTGAGGAGG
CAGCCCTGACCCGCAGGATCTGTAATCCTACTTCCCTGCTGGCTGCCTTTAGACATGGAGCTG
CTACACAGACAGGTCCTGGCTCTGCAGACACAGAGGGTCCTGCTGGAGAAGAGGGCGGAAGGC
TTCAGCCTGGCAACGGAACCTGGGCTACCCCCCTGGCTATGCTGTGCTTGCTGGTGCTGACGG
GCCTGTCTGTGCTCATTGTGGCCATCCACATCCTGGAGCTGCTCATCGATGAGGCTGCCATG
CCCCGAGGCATGCAGGGTACCTCCTTAGGCCAGGTCTCCTTCTCCAAGCTGGGCTCCTTTGG
TGCCGTCATTAGGTTGTACTCATCTTTTACCTAATGGTGTCTCAGTTGTGGGCTTCTATA
GCTCTCCACTCTTCCGGAGCCTGCGGCCCAGATGGCACGACACTGCCATGACGCAGATAATT
GGGAACTGTGTCTGTCTCCTGGTCCTAAGCTCAGCACTTCCTGTCTTCTCTCGAACCCCTGGG
GCTCACTCGCTTTGACCTGCTGGGTGACTTTGGACGCTTCAACTGGCTGGGCAATTTCTACA
TTGTGTTCTCTACAACGCAGCCTTTGCAGGCCTCACCACACTCTGTCTGGTGAAGACCTTC
ACTGCAGCTGTGCGGGCAGAGCTGATCCGGGCCTTTGGGCTGGACAGACTGCCGCTGCCCGT
CTCCGGTTTCCCCCAGGCATCTAGGAAGACCCAGCACCAGTGACCTCCAGCTGGGGGTGGGA
AGGAAAAAACTGGACACTGCCATCTGCTGCCTAGGCCTGGAGGGAAGCCCAAGGCTACTTGG
ACCTCAGGACCTGGAATCTGAGAGGGTGGGTGGCAGAGGGGAGCAGAGCCATCTGCACTATT
GCATAATCTGAGCCAGAGTTTGGGACCAGGACCTCCTGCTTTTCCATACTTAACTGTGGCCT
CAGCATGGGGTAGGGCTGGGTGACTGGGTCTAGCCCCTGATCCCAAATCTGTTTACACATCA
ATCTGCCTCACTGCTGTTCTGGGCCATCCCCATAGCCATGTTTACATGATTTGATGTGCAAT
AGGGTGGGGTAGGGGCAGGGAAAGGACTGGGCCAGGGCAGGCTCGGGAGATAGATTGTCTCC
CTTGCCCTCTGGCCCAGCAGAGCCTAAGCACTGTGCTATCCTGGAGGGGCTTTGGACCACCTG
AAAGACCAAGGGGATAGGGAGGAGGAGGCTTCAGCCATCAGCAATAAAGTTGATCCCAGGGA
AAAAAA

79/330

FIGURE 79

MEAPDYEVLSVREQLFHERIRECIISTLLFATLYILCHIFLTRFKKPAEFTTVDDDEDATV NK
IALELCTFTLAIALGAVLLLPFSIISNEVLLSLPRNYIIQWLNGSLIHGLWNLVFLFPNLSL
IFLMPFAYFFTESEGFAGSRKGV LGRVYETVVMLMLLTLLVLGMVWVASAIVDKNKANRESL
YDFWEYYLPYLYSCISFLGVLLLLVCTPLGLARMFSVTGKLLVKPRILLEDLEEQLYCSAFEE
AALTRRICNPTSCWLPLDMELLHRQVLALQTQRV LLEKRRKASAWQRNLGYPLAMLCLLVLT
GLSVLIVAIHILELLIDEAAMP RGMQGTSLGQVSFSKLG SFGAVIQVVLIFYLMVSSVVG FY
SSPLFRSLRPRWHD TAMTQIIIGNCVCLLV LSSALPVFSRTLGLTRFDLLGDFGRFNWLG NFY
IVFLYNAAFAGLTTLCLVKTFTA AVRAELIRAFGLDRLPLPVSGFPQASRKTQH Q

80/330

FIGURE 80

GGCTGCCGAGGGAAGGCCCTTGGGTGGTCTTGGTTGCTTGGCGGCGGCGGNTTCNTCCCC
GCTCGTCCTCCCCGGGCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGAAGC
ACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGTGTA
TTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCCTGACCCGCTTC
AAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCG

81/330

FIGURE 81

GACCGACCTTAAAGAGTGGGAGCAAAGGGAGGACAGAGCCTTTTAAAACGAGGCGGTGGTGC
CTGCCCTTTAAGGGCGGGGCGTCCGGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTTTC
TGTCGCAGGCTGCGAGGAAAGGCCCCCTAGGCTGGGTCTGGTGCTTGGCGGCGGGCGGCTTCCT
CCCCGTTGTCNTCCCCGGGCCCCAGAGGCACCTCGGCTTCAGTCATGCTGAGCAGAGTATGGA
AGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAGT
GTATTATATCAACACTTCTGTTTGCAACACTGTACATCNTCTGCCACATCTTCCTGACCCGC
TTCAAGAAGCCTGCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGATTGC
GCTCGAGCTGTGCACCTTTACCCTGGCAATTGCCCTGGGTGCTGTCCTGCTCCTGCCCTTCT
CCATCATCAGCAATGAGGTGCTGCACTCCC

82/330

FIGURE 82

GATGTGCTCCTTGGAGCTGGTGTGCAGTGTCTGACTGTAAGATCAAGTCCAAACCTGTTTT
GGAATTGAGGAACTTCTCTTTTGATCTCAGCCCTTGGTGGTCCAGGTCTTC**ATG**CTGCTGT
GGGTGATATTACTGGTCCTGGCTCCTGTCAGTGGACAGTTTGCAAGGACACCCAGGCCCAT
ATTTTCCTCCAGCCTCCATGGACCACAGTCTTCCAAGGAGAGAGAGTGACCCTCACTTGCAA
GGGATTTTCGCTTCTACTCACCACAGAAAACAAAATGGTACCATCGGTACCTTGGGAAAGAAA
TACTAAGAGAAACCCCAGACAATATCCTTGAGGTTTCAAGGAATCTGGAGAGTACAGATGCCAG
GCCCAGGGCTCCCCCTCTCAGTAGCCCTGTGCACTTGGATTTTTCTTCAGAGATGGGATTTCC
TCATGCTGCCCAGGCTAATGTTGAACTCCTGGGCTCAAGTGATCTGCTCACCT**TAG**GCCTCTC
AAAGCGCTGGGATTACAGCTTCGCTGATCCTGCAAGCTCCACTTTCTGTGTTTGAAGGAGAC
TCTGTGGTTCTGAGGTGCCGGGCAAAGGCGGAAGTAACACTGAATAATACTATTTACAAGAA
TGATAATGTCCTGGCATTCTTAATAAAAGAACTGACTTCCAAAAAAAAAAAAAAAAAAAAA
AAA

83/330

FIGURE 83

MLLWVILLVLAPVSGQFARTPRPIIFLQPPWTTVFQGERVTLTCKGFRFYSPQKTKWYHRYL
GKEILRETPDNILEVQESGEYRCQAQGSPLSSPVHLD FSSEMGFPHAAQANVELLGSSDLLT

84/330

FIGURE 84

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCCCGGTGT
GAGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGCGGAGGAGGCTGTGAG
GAGTGTGTGGAACAGGACCCGGGACAGAGGAACC**ATG**GCTCCGCAGAACCTGAGCACCTTTT
GCCTGTTGCTGCTATACCTCATCGGGGCGGTGATTGCCGGACGAGATTTCTATAAGATCTTG
GGGGTGCCTCGAAGTGCCTCTATAAAGGATATTAAAAAGGCCTATAGGAACTAGCCCTGCA
GCTTCATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCCAGGATCTGGGTG
CTGCTTATGAGGTTCTGTGATAGTGAAGAACGGAACAGTACGATACTTATGGTGAAGAA
GGATTAAAAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTTCACACTTCTTTGGGGATTT
TGGTTTCATGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTCCAAGAGGAAGTGATA
TTATTGTAGATCTAGAAGTCACTTTGGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTT
AGAAACAAACCTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTCGGCAAGAGAT
GCGGACCACCCAGCTGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAAT
GCCCTAATGTCAAACCTAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTG
AGAGACGGCATGGAGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGG
AGATTTACGGTTCCGAATCAAAGTTGTCAAGCACCCAATATTTGAAAGGAGAGGAGATGATT
TGTACACAAATGTGACAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACT
CACTTGGATGGTCACAAGGTACATATTTCCCGGGATAAGATCACCAGGCCAGGAGCGAAGCT
ATGGAAGAAAGGGGAAGGGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAA
TCACTTTTGATGTGGATTTTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAA
CAGCTACTGAAACAAGGGTCAGTGCAGAAGGTATACAATGGACTGCAAGGATAT**TG**AGAGTG
AATAAAATTGGACTTTGTTTAAAATAAGTGAATAAGCGATATTTATTATCTGCAAGGTTTTT
TTGTGTGTGTTTTTGTTTTTATTTTCAATATGCAAGTTAGGCTTAATTTTTTTATCTAATGA
TCATCATGAAATGAATAAGAGGGCTTAAGAATTTGTCCATTTGCATTGCGAAAAGAATGACC
AGCAAAAGGTTTACTAATACCTCTCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGT
TTCAAGAATTAAAGCTGCAAGAGGACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGA
GTTGTTAGCAATTTCAATCAAAATGCCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTG
TTATTTTTTA

85/330

FIGURE 85

MAPQNLSTFCLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKLALQLHPDRNPDDPQ
AQEKFQDLGAAYEVLSDSEKRRKQYDTYGEGLKDGHQSSHGDI FSHFFGDFGFMFGGT PRQQ
DRNIPRGSDIIVDLEVTLEEYAGNFVEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQ
MTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGM EYPFFIGEGEPHVDGEPGDLRFRIKVVKH
PIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVVHISRDKITRPGAKLWKKGEGLPNFD
NNNIKGSLLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

Important features:**Signal peptide:**

amino acids 1-22

Cell attachment sequence.

amino acids 254-257

Nt-dnaJ domain signature.

amino acids 67-87

Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

N-glycosylation site.

amino acids 5-9, 261-265

Tyrosine kinase phosphorylation site.

amino acids 253-260

N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

Amidation site.

amino acids 164-168

86/330

FIGURE 86

TGGGACCAGGGAACCCCGGGCCCCCGGTGGAGNGCCTAACAGGCCGGTGGNTGCGACCGAA
GCGGCGGGCGGAGGAGGTTTTGAGGATTTTTGGAACAGGACCCGGACAGAGGAACCATGGTT
CCGCAGAACNTGAGCACNTTTTGCCTGTTGNTGNTATACTTCATCGGGGCGGTGATTGCCGG
ACGAGATTTNTATAAGATTTTGGGGTGCCTNGAAGTGCCTTNTATAAAGGATATTAAAAAGG
CCTATAGGAAACTAGCCCTGCAGNTTTATCCCGACCGGAACCCTGATGATCCACAAGCCCAG
GAGAAATTCCAGGATTTGGGTGCTGCTTATGAGGTTNTGTCAGATAGTGAGAAACGGAAACA
GTACGATAATTATGGTGAAGAAGGATTAAAAGATGGTNATCAGAGCTCCCATGGAGACATTT
TTTCACACTTNTTTGGGGATTTTGGTTTCATGTTTGGAGGAACCCCTNGTCAGCAAGACAGA
AATATTCCAAGAG

87/330

FIGURE 87

GGCACGAGGCGGCGGGGCAGTCGCGGGGATGCGCCCGGGAGCCACAGCCTGAGGCCCTCAGGT
CTCTGCAGGTGTCGTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCA
GCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACTGAGTCAGGAGCCCTCTGGAAGC**ATGG**
AGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCATCTTTCTGGCTTCGTTTGCAGCC
TTGGTGCTGGTTTGCAGGCAGCGCTACTGCCGGCCGCGAGACCTGCTGCAGCGCTATGATTC
TAAGCCCATTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAC
TGGACGATGTCGTTATCACCAACCCCCACATTGAGGCCATTCTGGAGAATGAAGACTGGATC
GAAGATGCCTCGGGTCTCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGAC
AGAGAAGCTTGTTGCCATGACAATGGGCTCTGGGGCCAAGATGAAGACTTCAGCCAGTGTCA
GCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGGATGATGTTGTGAAGTCGATG
TACCCTCCGTTGGACCCCCAACTCCTGGACGCACGGACGACTGCCCTGCTCCTGTCTGTCAG
TCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACC
AGTCTCTGTGCGCTGCTGAGGAGCATTTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAG
CCAGATAAAGGCCTCCCAGGCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATTT**TAGT**GCCT
ACAGGCCAGCAGCTAGCCATGAAGGCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTT
CTACTTTTTCTATAGAGTTAGTTGTTCTCCACGGCTGGAGAGTTCAGCTGTGTGTGCATAG
TAAAGCAGGAGATCCCCGTCAGTTTATGCCTCTTTTGCAGTTGCAAACGTGGCTGGTGAGT
GGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACTCTCTGCAAGAGGAGTATTGAAAA
CTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCT
AAGAAATCAAGAGGTTTCACATTAAAATTAGAATTTCTGGCCTCTCTCGATCGGTCAGAATG
TGTGGCAATTCTGATCTGCATTTTCAGAAGAGGACAATCAATTGAACTAAGTAGGGGTTTC
TTCTTTTGGCAAGACTTGTA CTCTCACCTGGCCTGTTTCATTTATTTGTATTATCTGCCT
GGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTG CAGGTTTGGGTTTGAAGCTGAGGAACT
ACAAAGTTGATGATTTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTG
GATAGTAAATTTATACTTATGTTTCCCTCAAAAAAAAAAAAAA

88/330

FIGURE 88

METVVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPSEL
ELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKTSAS
VSDIIVVAKRISPRVDDVVKSMYPPLDPKLLDARTTALLSVSHLVLVTRNACHLTGGLDWI
DQSLSAEEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

89/330

FIGURE 89

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCC
CCACCACTGCAGCC**ATG**ATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGA
TTTGGAGTGTTTTTCTGTTCTTTGGAATGATTCTCTTTTTTGACAAAGCACTACTGGCTAT
TGGAATGTTTTATTTGTAGCCGGCTTGGCTTTTGTAATTGGTTTAGAAAGAACATTCAGAT
TCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTTTCTGGGTGGTGTATTTGTAGTC
CTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTTCTCTTGTTTCAG
GGGCTTCTTTCCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCTTGATCCCTCCTAAAT
TTACCTGGAATTAGATCATTTGTAGATAAAGTTGGAGAAAGCAACAATATGGTAT**TAA**CAACA
AGTGAATTTGAAGACTCATTTAAAATATTGTGTTATTTATAAAGTCATTTGAAGAATATTCA
GCACAAAATTAAATTACATGAAATAGCTTGTAATGTTCTTTACAGGAGTTTAAACGTATAG
CCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTACTCAAGTGA
ACTAAGAAGAAGTCAGCAAGCAAACCTGAGAGAGGTGAAATCCATGTTAATGATGCTTAAGAA
ACTCTTGAAGGCTATTTGTGTTGTTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAA
CTGTGGTGCCTGTTTCTTTTCTTTTTATTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCT
TTTTAGAAGTGTCCACTGCAATGGCAAAAATATTTCCAGTTGCACTGTATCTCTGGAAGTGA
TGCATGAATTCGATTGGATTGTGTCATTTTAAAGTATTAAACCAAGGAAACCCCAATTTTG
ATGTATGGATTACTTTTTTTTGNCGNCAGGGCC

90/330

FIGURE 90

MISLTDTQKIGMGLTGFGVFFLFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQK
HKMKATGFFLGGVFVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFI RRVPVLGSLLNLPGI
RSFVDKVGESNNMV

Important features:

Transmembrane domains:

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

N-myristoylation sites.

amino acids 11-16, 51-56 and 116-121

Aminoacyl-transfer RNA synthetases class-II protein.

amino acids 49-59

91/330

FIGURE 91

GAAGACGTGGCGGCTCTCGCCTGGGCTGTTTCCCGGCTTCATTTCTCCCGACTCAGCTTCCC
ACCNTGGGCTTTCCGAGGTGCTTTCGCCGCTGTCCCCACCACTGCAGCCATGATCTCCTTAA
CGGACACGCAGAAAATTGGAATGGGATTAACCGGATTTGGAGTGTTTTTCCTGTTCTTTGGA
ATGATTCTCTTTTTTTGACAAAGCACTACTGGCTATTGGAAATGTTTTATTTGTAGCCGGCTT
GGCTTTTGTAATTGGTTTAGAAAGAACATTCAGATTCTTCTTCCAAAAACATAAAATGAAAG
CTACAGGTTTTTTTTCTGGGTGGTGTATTTGTAGTCCTTATTGGTTGGCCTTTGATAGGCATG
ATCTTCGAAATTTATGGATTTTTTCTCTTGTTT

92/330

FIGURE 92

GGCACGAGGCTGAACCCAGCCGGCTCCATCTCAGCTTCTGGTTTCTAAGTCCATGTGCCAAA
GGCTGCCAGGAAGGAGACGCCCTTCCTGAGTCCTGGATCTTTCTTCCTTCTGGAAATCTTTGA
CTGTGGGTAGTTATTTATTTCTGAATAAGAGCGTCCACGCATC**CATG**GACCTCGCGGGACTGC
TGAAGTCTCAGTTCCTGTGCCACCTGGTCTTCTGCTACGTCTTTATTGCCTCAGGGCTAATC
ATCAACACCATTTCAGCTCTTCACTCTCCTCCTCTGGCCCATTAACAAGCAGCTCTTCCGGAA
GATCAACTGCAGACTGTCCTATTGCATCTCAAGCCAGCTGGTGATGCTGCTGGAGTGGTGGT
CGGGCACGGAATGCACCATCTTCACGGACCCGCGCGCCTACCTCAAGTATGGGAAGGAAAAT
GCCATCGTGGTTCTCAACCACAAGTTTGAAATTGACTTTCTGTGTGGCTGGAGCCTGTCCGA
ACGCTTTGGGCTGTTAGGGGGCTCCAAGGTCCTGGCCAAGAAAGAGCTGGCCTATGTCCCAA
TTATCGGCTGGATGTGGTACTTCACCGAGATGGTCTTCTGTTTCGCGCAAGTGGGAGCAGGAT
CGCAAGACGGTTGCCACCAGTTTGCAGCACCTCCGGGACTACCCCGAGAAGTATTTTTTTCCT
GATTCACTGTGAGGGCACACGGTTCACGGAGAAGAAGCATGAGATCAGCATGCAGGTGGCCC
GGGCCAAGGGGCTGCCTCGCCTCAAGCATCACCTGTTGCCACGAACCAAGGGCTTCGCCATC
ACCGTGAGGAGCTTGAGAAATGTAGTTTCAGCTGTATATGACTGTACACTCAATTTCAGAAA
TAATGAAAATCCAACACTGCTGGGAGTCCTAAACGGAAAGAAATACCATGCAGATTTGTATG
TTAGGAGGATCCCCTGGAAGACATCCCTGAAGACGATGACGAGTGCTCGGCCTGGCTGCAC
AAGCTCTACCAGGAGAAGGATGCCTTTCAGGAGGAGTACTACAGGACGGGCACCTTCCGAGA
GACGCCCATGGTGCCCCCCCCGGCGGCCCTGGACCCTCGTGAACTGGCTGTTTTGGGCCTCGC
TGGTGCTCTACCCTTTCTTCCAGTTCCTGGTACGATGATCAGGAGCGGGTCTTCCCTGACG
CTGGCCAGCTTCATCCTCGTCTTCTTTGTGGCCTCCGTGGGAGTTCGATGGATGATTGGTGT
GACGGAAATTGACAAGGGCTCTGCCTACGGCAACTCTGACAGCAAGCAGAACTGAATGACT
GACTCAGGGAGGTGTCACCATCCGAAGGGAACCTTGGGGAACCTGGTGGCCTCTGCATATCCT
CCTTAGTGGGACACGGTGACAAAGGCTGGGTGAGCCCCTGCTGGGCACGGCGGAAGTCACGA
CCTCTCCAGCCAGGGAGTCTGGTCTCAAGGCCGGATGGGGAGGAAGATGTTTTGTAATCTTT
TTTTCCCCATGTGCTTTAGTGGGCTTTGGTTTTCTTTTTGTGCGAGTGTGTGTGAGAATGGC
TGTGTGGTGAGTGTGAACTTTGTTCTGTGATCATAGAAAGGTATTTTAGGCTGCAGGGGAG
GGCAGGGCTGGGGACCGAAGGGGACAAGTTCCCCTTTCATCCTTTGGTGCTGAGTTTTCTGT
AACCCTTGGTTGCCAGAGATAAAGTGAAAAGTGCTTTAGGTGAGATGACTAAATTATGCCTC
CAAGAAAAAAAATTAAGTGCTTTTCTGGGTCAAAAAAAAAAAAA

93/330

FIGURE 93

MDLAGLLKSQFLCHLVFCYVFIA SGLIINTIQLFTLLLWPINKQLFRKINCRLSYCISSQLV
MLEWWSGTECTIFTDPRAYLKYGKENAIVVLNHNKFEIDFLCGWSLSERFGLLGSKVLAKK
ELAYVPIIGWMWYFTEMVFCSRKWEQDRKTVATSLQHLDYPEKYFFLIHCEGTRFTEKKHE
ISMQVARAKGLPRLKHHLLPRTKGFAITVRSLRNVVSAVYDCTLNFRNNENPTLLGVLNGKK
YHADLYVRRIPLEDIPEDDDECSAWLHKLYQEKDAFQEEYYRTGTFFPETPMVPPRRPWTLVN
WLFWASLVLYPFFQFLVSMIRSGSSLTLASFILVFFVASVGVRWMIGVTEIDKGSAYGNSDS
KQKLND

94/330

FIGURE 94

CTGAGGCGGCGGTAGC**ATG**GAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCG
GCGCACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAA
GTAAAAGGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTA
TACAATTGACATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAG
GCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGT
TGGTACAAATTCCGTCGTCATTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAA
AACTTGCAGGAGCATTTTTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAA
TAACAGAAAGCTGCTCTACTCATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTT
TTTCACAGGGTACCTTTAGTGTTGCCAATCTGGGCATGTCTGAACAACTGGGTTATAAAAC
TGTATCAGGTTCCGTGTATGTCCACTGGTTTTAGCCGAGCAGTACAAACACACAGCTCTAAAT
TTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGCTTCATTA
CAAGAGGAATTAAAGAGTATATGCAAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACT
AGTAAAGGATGTAAACAGATTAAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAGG
CAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGAACATTTTTCTTTGTCAGGCATTA
CGGACCTTTTTTCCAAATCTGAATTTCTTCATTCATGTGTTATGTCTTTAAAAAATAGACA
TGTTTCTAAAAGTAGCTGTAACCTACAACCACCATCTCGATGTAGTAGACAATCTGACCTTAA
TGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTAAGCAT
AAAGCCTTAGACTTAGATGACAGATGGCAATTCAGAGATCTCGGTTGTTAGATACACAAGA
CAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCC
CAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACA
TTTT**TGA**TCCTTTTAAACCTTACAAGGAGATTTTTTTTATTTGGCTGATGGGTAAAGCCAAACAT
TTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTAC
CTGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAAC
ATCAGATGCTTTTTATTTCCAAACCTTTTTTTTACCTTTCACTAAGTTGTTGAGGGGAAGGCT
TACACAGACACATTCCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAA
TCCCAGCACTTAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCC
TGGGCAACGTATTGAGACCATGTCTATTAATAAATAAATGGAAGCAAGAATAGCCTTAT
TTTCAAAATATGGAAAGAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCTCCTTAAG
TGATACTTTTTTGAAGTACATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCA
ATAAATTTGCAAAACATCATCTAAAATTTAAAAAAAAAAAAAAAAAAAAAAAAA

95/330

FIGURE 95

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTIDIQ
KYIPCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEH
FSNQDLVFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSC
MSTGFSRAVQTHSSKFFEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVN
RLKREIEKRRGAQIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSS
CNYNHHLDVVDNLTLMVEHTDIPEASPASTPQIIKHKALDLDDRWFKRSRLLDTDKRSKA
NTGSSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTF

96/330

FIGURE 96

GGCACAGCCGCGCGGGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGC
CCAAGCAGCGCGCAGCGAACGCCCCGCCGCCGCCACACCCTCTGCGGTCCCCGCGGCGCCTGCCACCCCTTCCCT
CCTTCCCCGCGTCCCCGCGCTCGCCGGCCAGTCAGCTTGCCGGGTTGCTGCCCCGCGAAACCCCGAGGTACCA
GCCCCGCGCTCTGCTTCCCTGGGCGCGCGCGCCGCTCCACGCCCTCCTTCTCCCTGGCCCCGGCGCCTGGCACC
GGGGACCGTTGCTGACGCGAGGCCAGCTCTACTTTTCGCCCCGCGTCTCCTCCGCTGCTCGCCTCTTCCAC
CAACTCCAACCTCTTCTCCCTCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCCCGCTGCCGTAG
CGCCGCTTCCCGTCCGGTCCCAAAGGTGGGAACGCGTCCGCCCGCGCCCGCACC**ATG**GCACGGTTCGGCTTGCC
CGCGCTTCTCTGCACCCTGGCAGTGCTCAGCGCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGG
AAGTGGCAGCTCTTTACGTGTCCAAAGGCTTCAACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCAT
TTGAAGATCTGTCCCAGGGTTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGA
TGATTTCAAAGTGTGGTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTTGCTTCACGTTACAAGAAGTTTG
ATGAATTTCTCAAAGAACTACTTGAAAATGCAGAGAAATCCCTGAATGATATGTTTGTGAAGACATATGGCCAT
TTATACATGCAAATTTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGT
GAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATGTTCCGCTGGTGAACCTCCAGT
ACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCT
CGCAAATTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCGTACTTTGCTCAAGGCTTAGCGGTTGCGGG
AGATGTGCTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAGATGATCTACT
GCTCCCACTGCCGGGGTCTCGTGACTGTGAAGCCATGTTACAACCTACTGCTCAAACATCATGAGAGGCTGTTTG
GCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGA
GGGTCCTTTCAACATTGAATCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGG
ATAATAGTGTCAAGTGTCTCAGAAGGTTTCCAGGGATGTGGACCCCCCAAGCCCTCCAGCTGGACGAATT
TCTCGTTCCATCTCTGAAAGTGCCTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCAACACAGC
AGCTGGCACTAGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTTCTGGTCCT
CCCTTCCGAGCAACGTTTGAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGG
AAAGGCAAAAGCAGGTACCTGTTTGCAGTGACAGGAAATGGATTAGCCAACCAGGGCAACAACCCAGAGGTCCA
GGTTGACACCAGCAAACCAGACATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGA
AGAATGCATACAATGGGAACGACGTGGACTTCTTTGATATCAGTGATGAAAGTAGTGGAAGGAAGTGGAAAGT
GGCTGTGAGTATCAGCAGTGCCCTTCAGAGTTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGA
GAAAGCCGACAGTGCTGGTGTCCGTCTGGGGCACAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGG
TTATGCAGAGAGAGTGGAGAT**TAA**TTCTCAAACCTCTGAGAAAAAGTGTTATCAAAAAAGTTAAAAGGCACCAAGTT
ATCACTTTTCTACCATCCTAGTGACTTTGCTTTTTTAAATGAATGGACAACAATGTACAGTTTTTACTATGTGGC
CACTGGTTTAAGAAGTGCTGACTTTGTTTTCTCATTCAAGTTTTGGGAGGAAAAAGGGACTGTGCATTGAGTTGGT
TCCTGCTCCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAACTATAGTTAGTTGTGCATTTGTGA
TTTTATCACTCTATTATTTGTTTGTATGTTTTTTTCTCATTTCGTTTGTGGGTTTTTTTTTCCAACTGTGATCT
CGCCTTGTTTCTTACAAGCAAACCAGGGTCCCTTCTTGGCACGTAACATGTACGTATTTCTGAAATATTAATA
GCTGTACAGAAGCAGGTTTTATTTATCATGTTATCTTATTAAGAAAAAGCCCCAAAAAGC

97/330

FIGURE 97

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQ
GSTCCSQEMEKEYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMF
VKTYGHLQNSSELFKDLFVELKRYVVGNNLEEMLNDFWARLLERMFLVNSQYHFTDEY
LECVSKYTEQLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHAL
LKMIYCSHCRGLVTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIES
VMDPIDVKISDAIMNMQDNSVQVSQKVFQGC GPPKPLPAGRISRSISESAFSARFRPHHPEE
RPTTAAGTSLDRLVTDVKEKLKQAKKFWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLE
AVTGNGLANQGNNPEVQVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGE
GSGSGCEYQQCPSEFDYNATDHAGKSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

98/330

FIGURE 98

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCT
GACCTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAA
GCAACTTACAGCTGCACCGACAGTTGCG**ATG**AAGTTCTAATCTCTTCCCTCCTCCTGTTGC
TGCCACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTGCGCCAGAGGCCAC
AGGGACCGAGGCCAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAA
AGATTGGTTCCTGAGAGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGC
AGTGCCCTGTGATCATTTCAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGA
AAGCCAAACAAGCATTCCAGAGCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTT
TGCTCTGCCTTTG**TAG**GAGCTCTGAGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAA
GACAGTGAGCACACCTACCAGACACTCTTCTTCTCCACCTCACTCTCCCACTGTACCCACC
CCTAAATCATTCAGTGCTCTCAAAAAGCATGTTTTTCAAGATCATTTTGTTTGTTGCTCTC
TCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCCTTACCCAGGCTTAGGCTT
AATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCTAGCTAGTGTCATTTAACCTTAAATGC
AATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTTAAATGTCAAAAAAAAAAAAAAAAAA

99/330

FIGURE 99

MKVLISLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRR
KFMTVSGLPKKQCPDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

100/330

FIGURE 100

A**ATG**GCTGTCTTAGTACTTCGCCTGACAGTTGTCCTGGGACTGCTTGTCTTATTCCTGACCT
GCTATGCAGACGACAAACCAGACAAGCCAGACGACAAGCCAGACGACTCGGGCAAAGACCCA
AAGCCAGACTTCCCCAAATTCCTAAGCCTCCTGGGCACAGAGATCATTGAGAATGCAGTCGA
GTTTCATCCTCCGCTCCATGTCCAGGAGCACAGGATTTATGGAATTTGATGATAATGAAGGAA
AACATTCATCAAAG**TGA**CATCCTCAGGACACACCCATGTGGCTCCTGGACAATCCAAGAGCA
GCCAAATCCTGCTTTTCCAGTTTGGCTCCACAAGTCCTCCAGGACAGAGCCCTCAAAGCAAC
TCCCAACGAGTTCTCAGGATTCAGGCTCTGGCTTCAACCAAACAGAACTCATTTTGAACACC
CTGACTGCATTTTTGCTTTTAGAAAGTTAGAATAAATATGGCGCTTTGGGATCACATAGTTG
ATGGAGAGGAAA

101/330

FIGURE 101

MAVLVLRLTVVLGLLVFLTCYADDKPDKPDDKPDDSGKDPKPDFPKFLSLLGTEIIENAVE
FILRSMRSTGFMEFDDNEGKHSSK

102/330

FIGURE 102

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCT
CAGAGCTGGTCTGCC**ATG**GACATCCTGGTCCCACCTCCTGCAGCTGCTGGTGCTGCTTCTTAC
CCTGCCCCCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCCTGTGCAAAAGCTACTTCC
CCTACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAG
CTCTTCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGG
CTGCGGAACCGGAGCCAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACC
CAAATCCCCACTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATAT
GAGCGGTTTGTGGTGGCTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGT
GGTGGTCTGCACTCTGGTGCTGTGCTCTGTGCAGAGCCCCAAGGAAGGTCCTGCAGGAGGTCC
GGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGCATGTGGCAGAACCATATGGA
AGCTGGGCCTTCATGTGGCAGCAAGTTTTTCGAGCCCACCTGGAAACACATTGGGGATGGCTG
CTGCCTCACCAGAGAGACCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGG
AACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTC
AAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCCTTCCCCAGCCTCCAATTAGAACA
AGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCT**TAG**CAGAATGAGAGAAGACATT
CATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGCAATCTCTAACTTCAATC
CCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGGACCC
TGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTC
CCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACACCCATGCGT
CTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCT
CTCCCCACTACCACCTTCTTCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGGAT
GCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGA
AACCACG

103/330

FIGURE 103

MDILVPLLQLLVLLLTLP LHL MALLGCWQPLCKSYFPYLM AVLTPKSNRKMESKKRELF S Q I
KGLTGASGKVALLELGCGTGANFQFYPPGCRVTCLDPNPHFEKFLTKSMAENRHLQYERFVV
APGEDMRQLADGSMDVVVCTLVLC SVQSPRKVLQEVRRLRPGGVLFFWEHVAEPYGSWAFM
WQQVFEPTWKHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPVGPHIMGKAVKQSFP
SSKALICSFPSLQLEQATHQPIYLPLRGT

104/330

FIGURE 104

GTGGGATTTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGAAAGTTGCCTCATCGCAGG
CAGATGTTGGGGCTTTGTCCGAACAGCTCCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAA
ACTAATATTTATATGACAGAAGAAAAAG**ATG**CATTCCGTAAAGTAAACATCATCATCTTGG
TCCTGGCTGTTGCTCTCTTCTTACTGGTTTTGCACCATAACTTCCTCAGCTTGAGCAGTTTG
TTAAGGAATGAGGTTACAGATTGAGGAATTGTAGGGCCTCAACCTATAGACTTTGTCCCAA
TGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCTGTGGTCATCGCTGCATCTG
AAGACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTGAGCACAACACTCGCTCCAAT
GTGATTTTCTACATTGTTACTCTCAACAATACAGCAGACCATCTCCGGTCTGGCTCAACAG
TGATTCCCTGAAAAGCATCAGATACAAAATTGTCAATTTTGACCCTAAACTTTTGGAAGGAA
AAGTAAAGGAGGATCCTGACCAGGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTAC
TTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCA
AGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAG
AAGATTGTGATTGAGCCTCTACTAAAGTTGTCATCCGTGGAGCAGGAAACAGTACAATTAC
ATTGGCTATCTTGACTATAAAAAGGAAAGAATTTCGTAAGCTTTCCATGAAAGCCAGCACTTG
CTCATTTAATCCTGGAGTTTTTTGTTGCAAACCTGACGGAATGGAAACGACAGAATATAACTA
ACCAACTGGAAAAATGGATGAAACTCAATGTAGAAGAGGGACTGTATAGCAGAACCTGGCT
GGTAGCATCACAAACACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCGATCC
TATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTA
AGGCTGCCAAGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATAT
ACTGATGTTTGGGAAAAATGGTATATTCCAGACCCAACAGGCAAATTCAACCTAATCCGAAG
ATATACCGAGATCTCAAACATAAAG**TGA**AACAGAATTTGAACTGTAAGCAAGCATTCTCAG
GAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTCAATGCCTATCGGTA
GCAAGCCATGGAAAAAGATGTGTCAGCTAGGTAAAGATGACAACTGCCCTGTCTGGCAGTC
AGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTTTTCTT
ACTACAATGCTGAATGACTGGAAAGAAGAACTGATATGGCTAGTTCAGCTAGCTGGTACAGA
TAATTCAAACTGCTGTTGGTTTTAATTTTGTAACCTGTGGCCTGATCTGTAAATAAACTT
ACATTTTTC

105/330

FIGURE 105

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTDSGIVGPQPIDFVPNALRHAVDGR
QEEIPVVIAASEDRLGGAIAAINSIQHNTRSNVIFYIVTLNNTADHLRSWLNSDSLKSIRYK
IVNFDPKLLEGKVKEDPDQGESMKPLTFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNT
ALKPGHAAAFSEDCDSASTKVVIRGAGNQYNYIGYLDYKKERIRKLSMKASTCSFNPGVFVA
NLTEWKRQNITNQLEKWMKLNVEEGLYSRTLGSITTPPLLIVFYQQHSTIDPMWNVRLHLS
SAGKRYSPQFVKAALLHWNGHLKPWGRTASYTDVWEKWYIPDPTGKFNLIRRYTEISNIK

106/330

FIGURE 106

TGGTTTTTGGCCCATAAATTCCTCAGCTTGAGCAGTTTGTTAAGGAATGAGGTTACAGATT
CAGGAATTNTAGGNCCTCAACCTNTAGANTTTGTCCCAAATGTTCTCCGACATGCAGTAGAT
GGGAGACAAGAGGAGATTCTGTGGTCATCGCTGCATNTGAAGACAGGCTTGGGGGGGCCAT
TGCAGCTATAAACAGCATTCAGCACAACTCGNTCCAATGTGATTTTCTACATTGTTACTC
TCAACAATACAGCAGACCATNTCCGGTCCTGGNTCAACAGTGATTCCCTGAAAAGCATCAGA
TACAAAATTGTCAATTTTGACCCTAACTTTTGGAAGGAAAAGTAAAGGAGGATCCTGACCA
GGGGGAATCCATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGTTCCCAGCG
CAAAGAAGGCCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTAC
AATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATTGTGATTCAGCCTCTAC
TAAAGTTGTCATCCGTGGAGCAGGAAA

107/330

FIGURE 107

CGACGCTCTAGCGGTTACCGCTGCGGGCTGGCTGGGCGTAGTGGGGCTGCGCGGCTGCCACG
GAGCTAGAGGGCAAGTGTGCTCGGCCCAGCGTGCAGGGAACGCGGGCGGCCAGACAACGGGC
TGGGCTCCGGGGCCTGCGGCGCGGGCGCTGAGCTGGCAGGGCGGGTCGGGGCGCGGGCTGCA
TCCGCATCTCCTCCATCGCCTGCAGTAAGGGCGGCCGCGGCGAGCCTTTGAGGGGAACGACT
TGTCGGAGCCCTAACCAGGGGTGTCTCTGAGCCTGGTGGGATCCCCGGAGCGTCACATCACT
TTCCGATCACTTCAAAGTGGTTAAAACTAATATTTATATGACAGAAGAAAAAGATGTCATT
CCGTAAAGTAAACATCATCATCTTGGTCTGGGCTGTTGCTCTCTTCTTACTGGTTTTGCAC
CATAACTTCCTCAGCTTGAGGCAGTTTGTTAAGGAATGAGGTTACAGATTCAGGAATTGTAG
GGCCTCAACCTATAGGACTTTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGA
GGAGATTCCTGTGGTCATCGCTGCATCTGAAGACAGGCTTGGGGGGGCCATTGCAGCTATAA
ACAGCATTCAGCACAACACTCGCTCCAATGTGATTTTCTACATTGTTACTCTCAACAATACA
GCAGACCATCTCCGGTCCTGGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAATTG
TCAATTTTGACCCTAACTTTTGAAGGAAAAGTAAAGGAGGATCCTGACCAGGGGGAATCC
ATGAAACCTTTAACCTTTGCAAGGTTCTACTTGCCAATTCTGGGTTCCCAGCGCAAAGAAGG
CCATATACATGGATGATGATGTAATTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCA
CTGAAGCCAGGACATGCAGCTGCATTTTTCAGAAGATTGTGATTCAGCCTCTACTAAAGTTGT
CATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATAAAAAGGAAAGAA
TTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTGCAAAC
CTGACGGAATGGAAACGACAGAATATACTAACCAACTGGAAAAATGGATGAACTCAATGT
AGAAGAGGGACTGTATAGCAGAACCCTGGCTGGTAGCATCACAAACACCTCCTCTGCTTATCG
TATTTTATCAACAGCACTCTACCATCGATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGT
GCTGGAAAACGATATTCACCTCAGTTTGTAAAGGCTGCCAAGTTACTCCATTGGAATGGACA
TTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGGAAAAATGGTATATTCCA
GACCCAACAGGCAAATTCAACCTAATCCGAAGATATACCGAGATCTCAAACATAAAGTGAAA
CAGAATTTGAACTGTAAGCAAGCATTTCTCAGGAAGTCCTGGAAGATAGCATGCGTGGGAAG
TAACAGTTGCTAGGCTTCAATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAG
GTAAAGATGACAACTGCCCTGTCTGGCAGTCAGCTTCCCAGACAGACTATAGACTATAAAT
ATGTCTCCATCTGCCTTACCAAGTGTTTTCTTACTACAATGCTGAATGACTGGAAAGAAGAA
CTGATATGGCTAGTTCAGCTAGCTGGTACAGATAATTCAAAACCTGCTGTTGGTTTTAATTTT
GTAACCTGTGGCCTGATCTGTAAATAAACTTACATTTTTTCAATAGGTAAAAA

108/330

FIGURE 108

CTGCAGGTAGACATCTCCACTGCCCAGGAATCACTGAGCGTGCAGACAGCACAGCCTCCTCT
GAAGGCCGGCCATACCAGAGTCCTGCCTCGGCATGGGCCTCACCATTGAGGCAGCTCCACTG
TCTGTGCTGGTCTGAGGGTGCTGCCTGTC**ATG**GGGGGCAGCCATCTCCCAGGGGGCCCTCATC
GCCATCGTCTGCAACGGTCTCGTGGGCTTCTTGCTGCTGCTGCTCTGGGTCATCCTCTGCTG
GGCCTGCCATTCTCGTCTGCCGACGTTGACTCTCTCTCTGAATCCAGTCCCAACTCCAGCCC
TGGCCCCTGTCCTGAGAAGGCCCCACCACCCAGAAAGCCCAGCCATGAAGGCAGCTACCTGC
TGCAGCCCTGAAGGCCCCCTGGCCTAGCCTGGAGCCCAGGACC**TAA**GTCCACCTCACCTAGAG
CCTGGAATTAGGATCCCAGAGTTCAGCCAGCCTGGGGTCCAGAACTCAAGAGTCCGCCTGCT
TGGAGCTGGACCCAGCGGCCCAGAGTCTAGCCAGCTTGGCTCCAATAGGAGCTCAGTGGCCC
TAAGGAGATGGGCCTGGGGTGGGGGCTTATGAGTTGGTGCTAGAGCCAGGGCCATCTGGACT
ATGCTCCATCCCAAGGGCCAAGGGTCAGGGGCGGGTCCACTCTTTCCCTAGGCTGAGCACC
TCTAGGCCCTCTAGGTTGGGGAAGCAAACCTGGAACCCATGGCAATAATAGGAGGGTGTCCAG
GCTGGGCCCCCTCCCCTGGTCCTCCCAGTGTTTGCTGGATAATAAATGGAACCTATGGCTCTAA
AAAAAAAAAAAAAAAAAAAA

109/330

FIGURE 109

MGA AISQ GALIA IVCNGLVG FLLLLLWVILCWACHSR LPTLTSLNPVPTPALAPVLRRPHH
PRSPAMKAATCCSPEGPWPSLEPRT

110/330

FIGURE 110

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCA
GTTCCCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTA
CTCCCTATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGA
TC**ATG**TCGGGAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCC
ATGATGTTTACCTTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTCAATTGGTTAT
TTTGGGATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACC
TCAGCATAGAATTGGACACAGAAAGGGAAAATATGAAGTGCGTGCTGGGGTTTGCTATCGTA
TCCACAGGCATCACGGCAGTGCTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATT
GACAGTTGAGCTTTTCCAAATCACAAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCC
AGCCACTGTGGACATTTGCCATCCTCATTTTCTTCTGGGTCCCTCTGGGTGGCTGTGCTGCTG
AGCCTGGGAAGTGCAGGAGCTGCCCAGGTTATGGAAGGCGGCCAAGTGGAATATAAGCCCCT
TTCGGGCATTTCGGTACATGTGGTTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCA
TCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAGTGGTTACTTGTTATTTCAACAGAAGT
AAAAATGATCCTCCTGATCATCCATCCTTTCTGTCTCTCTCCATTCTCTTCTTCTACCATCA
AGGAACCGTTGTGAAAGGGTCATTTTTTAATCTCTGTGGTGAGGATTCCGAGAATCATTGTCA
TGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGTCCAGGTACCTGTTCCGA
TGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATGCATA
TACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAAATCT
TGTCCAAGAAGTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAATTTTTCTA
GGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTACAATCG
GGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTTGCCTACTTAGTAGCCC
ATAGTTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCCTGTGTTTTGCTGTTGAT
CTGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTT
CGTAAAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGA
ATGAGGAGGGAACAGAACTCCAGGCCATTGTGAGA**TAG**ATACCCATTTAGGTATCTGTACCT
GGAAAACATTTCCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTT
AGTGAATTTTTTTTTTAAAAGACCTAATAAACCTATTCTTCCTCAAAA

111/330

FIGURE 111

MSGRDTILGLCILALALSLAMMFTFRFITTTLLVHIFISLVILGLLFVCGVLWWLYDYTNDL
SIELDTERENMKCVLGFAIVSTGITAVLLVLI FVLRKRIKLTVELFQITNKAISSAPFLLFQ
PLWTFAILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFI
LACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVM
YMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFCTSAKDAFKIL
SKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAH
SFLSVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRN
EEGTELQAIVR

FIGURE 112

[illegible]

113/330

FIGURE 113

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKC
PAGCQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSL
SLPRWRESFIVLESKPKKGVTPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQ
LLAVTVAVATPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTTSSQNRPRADPGIQ
RQDPSGAAFQKPVGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLSFLIDGSTSIGKRRFR
IQKQLLADVAQALDIGPAGPLMGVVQYGDNPATHFNLKTHTNRDLKTAIEKITQRGGLSNV
GRAISFVTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAE
NEKQYVVEPNFANKAVCRTNGFYSLHVQSWFGLHKTQLPLVKRVCDTDRLACSKTCLNSADI
GFVIDGSSSVGTGNFRTVLQFVTNLTKEFEISDTRIGAVQYTYEQRLEFGFDKYSSKPD
LNAIKRVGYWSGGTSTGAAINFALEQLFKKSKPNKRKLMILITDGRSYDDVRI PAMAAHLKG
VITYAIGVAWAAQEELEVIATHPARDHSFFVDEFDNLHQYVPRIIQNICTEFNSQPRN

114/330

FIGURE 114

CAGGATGAACTGGTTGCAGTGGCTGCTGCTGCTGCGGGGGCGCTGAGAGGACACGAGCTCTA
TGCCTTTCCGGCTGCTCATCCCGCTCGGCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGT
GCGCCAGGTCCCGACGGCTCCGCGCCAGATCCCGCCCACTACAGTTTTTCTCTGACTCTAAT
TGATGCACTGGACACCTTGCTGATTTTGGGGAATGTCTCAGAATTCCAAAGAGTGGTTGAAG
TGCTCCAGGACAGCGTGGACTTTGATATTGATGTGAACGCCTCTGTGTTTGAAACAAACATT
CGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGGAAGTAGA
GGCTGGATGGCCCTGTTCCGGGCCTCTCCTGAGAATGGCTGAGGAGGCGGCCCGAAAACCTCC
TCCCAGCCTTTCAGACCCCCACTGGCATGCCATATGGAACAGTGAACCTTACTTCATGGCGTG
AACCCAGGAGAGACCCCTGTCACCTGTACGGCAGGGATTGGGACCTTCATTGTTGAATTTGC
CACCTGAGCAGCCTCACTGGTGACCCGGTGTTTGAAGATGTGGCCAGAGTGGCTTTGATGC
GCCTCTGGGAGAGCCGGTCAGATATCGGGCTGGTCGGCAACCACATTGATGTGCTCACTGGC
AAGTGGGTGGCCCAGGACGCAGGCATCGGGGCTGGCGTGGACTCCTACTTTGAGTACTTGGT
GAAAGGAGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTCTTAGAGTATAACAAAG
CCATCCGGAACCTACACCCGCTTCGATGACTGGTACCTGTGGGTTCAGATGTACAAGGGGACT
GTGTCCATGCCAGTCTTCCAGTCCTTGAGGGCCTACTGGCCTGGTCTTCAGAGCCTCATTGG
AGACATTGACAATGCCATGAGGACCTTCCTCAACTACTACACTGTATGGAAGCAGTTTGGGG
GGCTCCCGGAATTCTACAACATTCTCAGGGATACACAGTGGAGAAGCGAGAGGGCTACCCA
CTTCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACGGGGGATCCCACCCT
CCTAGAACTCGGAAGAGATGCTGTGGAATCCATTGAAAAAATCAGCAAGGTGGAGTGCGGAT
TTGCAACAATCAAAGATCTGCGAGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCCTG
GCCGAGACTGTGAAATACCTCTACCTCCTGTTTGACCCAACCAACTTCATCCACAACAATGG
GTCCACCTTCGACGCGGTGATCACCCCTATGGGGAGTGATCCTGGGGGCTGGGGGGTACA
TCTTCAACACAGAAGCTCACCCCATCGACCTTGCCGCCCTGCACTGCTGCCAGAGGCTGAAG
GAAGAGCAGTGGGAGGTGGAGGACTTGATGAGGGAATTCTACTCTCTCAAACGGAGCAGGTC
GAAATTTAGAAAAACACTGTTAGTTCGGGGCCATGGGAACCTCCAGCAAGGCCAGGAACAC
TCTTCTCACCAGAAAACCATGACCAGGCAAGGGAGAGGAAGCCTGCCAAACAGAAGGTCCCA
CTTCTCAGCTGCCCCAGTCAGCCCTTACCTCCAAGTTGGCATTACTGGGACAGGTTTTCT
AGACTCCTCA**TAA**CCACTGGATAATTTTTTTATTTTTATTTTTTTGAGGCTAAACTATAATA
AATTGCTTTTGGCTATCATAAAA

115/330

FIGURE 115

MPFRLLIPLGLLCALLPQHHGAPGPDGSAPDPAHYSFSLTLIDALDTLLILGNVSEFQRVVE
VLQDSVDFDIDVNASVFETNIRVVGGLLSAHLLSKKAGVEVEAGWPCSGPLLRMAEEAARKL
LPAFQTPTGMPYGTVNLLHGVNPGETPVTCTAGIGTFIVEFATLSSLTGDPVFEDVARVALM
RLWESRSDIGLVGNHIDVLTGKWVAQDAGIGAGVDSYFEYLVKGAILLQDKKLMAMFLEYNK
AIRNYTRFDDWYLWVQMYKGTVSMPVFQSLEAYWPGLQSLIGDIDNAMRTFLNYYTVWKQFG
GLPEFYNIPOGYTVEKREGYPLRPELIESAMYLYRATGDPTLLELGRDAVESIEKISKVECG
FATIKDLRDHKLDNRMESFFLAETVKYLYLLFDPTNFIHNNGSTFDAVITPYGECILGAGGY
IFNTEAHPIDLAALHCCQRLKEEQWEVEDLMREFYSLKRSRSKFQKNTVSSGPWEPPARPGT
LFSPENHDQARERKPAKQKVPLLSCPSQPFTSKLALLGQVFLDSS

116/330

FIGURE 116

AAAGTTACATTTTCTCTGGAACCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCCTGGGCCGGCTCTAGAACA
ATTCAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCCATGTTTGCTCACAGATG¹AAGTGGCCATTCTGCCTGCCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCAGTGATCGCGCCTGGAGA
AACAGTGTACTATTCTGTCTGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTGATGTCACTGATGACATC
ACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
CATTCGTGAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTTGCCCTTGTTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTTCGTCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCCGTGG
TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCCAGAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCTCCTGAGGAACCTCCTCAGGGCCTGGAT
CTCA**TAG**GTTTGCGGAAGGGCCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTTTCTGTTTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGGTTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACACAGAGTCTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAACTGGTGACACTCTACAGTCTGACTGATTCAG
TGTTTCTGGAGAGCAGGACATAAATGTATGATGAGAATGATCAAGGACTCTACACACTGGGT
GGCTTGGAGAGCCCACTTTCCCAGAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACTTCAAGCCCAATGCCGGTGACAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTCACAGCCCACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACTACTGTAAAGTGTTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTTC
TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
AAAAAAAAAA

117/330

FIGURE 117

MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLMLWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTDDITATVPYNLRVRATLGSQTS
SILKHPFNRNSTILTRPGMEITKDG FHLVIELEDLGPQFEFLVAYWRREPGAEEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPVLALFAFVGFM LILV
VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPEELLRAWIS

Important features:**Signal peptide:**

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation sites.

amino acids 40-43 and 134-137

Tissue factor proteins homology.

amino acids 92-119

Integrins alpha chain protein homology.

amino acids 232-262

118/330

FIGURE 118

TCCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAATTCAGGCTTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCCA
GTGATCGCGCCTGGAGAAACAGTGTACTATTCTGTCTGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTCAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTCCTTGTGGCCTANTGGAGGAGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCCAC
ATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

119/330

FIGURE 119

CGGACGCGTGGGCCGCCACCTCCGGAACAAGCC**ATG**GTGGCGGCGACGGTGGCAGCGGGCGTG
GCTGCTCCTGTGGGCTGCGGCCTGCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAGGCGG
TCAACATCCGGGGCAAACCTGGTGTGCTGGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTG
AATGTGGCCAGCGAGTGC GGCTTCACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCG
AGACCTGGGCCCCCACCACCTTTAACGTGCTCGCCTTCCCCTGCAACCAGTTTGGCCAACAGG
AGCCTGACAGCAACAAGGAGATTGAGAGCTTTGCCCCGCCGCACCTACAGTGTCTCATTCCCC
ATGTTTAGCAAGATTGCAGTCACCGGTACTGGTGCCCATCCTGCCTTCAAGTACCTGGCCCA
GACTTCTGGGAAGGAGCCCACCTGGAACCTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGG
TGGTAGGGGCTTGGGACCCAACCTGTGTCAGTGGAGGAGGTCAGACCCCAGATCACAGCGCTC
GTGAGGAAGCTCATCCTACTGAAGCGAGAAGACTT**TAA**CCACCGCGTCTCCTCCTCCACCA
CCTCATCCCGCCCACCTGTGTGGGGCTGACCAATGCAAACCTCAAATGGTGCTTCAAAGGGAG
AGACCCACTGACTCTCCTTCCTTTACTCTTATGCCATTGGTCCCATCATTCTTGTGGGGGAA
AAATTCTAGTATTTTGAATTATTTGAATCTTACAGCAACAAATAGGAACTCCTGGCCAATGAG
AGCTCTTGACCAGTGAATCACCAGCCGATACGAACGTCTTGCCAACAAAAATGTGTGGCAAA
TAGAAGTATATCAAGCAATAATCTCCCACCCAAGGCTTCTGTAAACTGGGACCAATGATTAC
CTCATAGGGCTGTTGTGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGC
CAAATAGGAGGCATTCAATGAACATTTTTTGCATATAAACCAAAAAATAACTTGTTATCAAT
AAAAACTTGCATCCAACATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTT
GTTATTTCTCTGTATTATTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCA
AACAATACCTCACGATATAAAATAAAAAATGAAAGTATCCTCCTCAAAAA

120/330

FIGURE 120

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVVNVASECGFTDQ
HYRALQQQLQRDLGPHHFNVLAFFPCNQFGQQEPDSNKEIESFARRTYSVSFFPMFSKIAVTGTG
AHPAFKYLAQTSKGKEPTWNFWKYLVPDGGKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

121/330

FIGURE 121

CGGACGCGTGGGCGGGCCGGGACGCGAGGGCAAAGCGAGCC**ATG**GCTGTCTACGTCGGGATGC
TGCGCCTGGGGAGGCTGTGCGCCGGGAGCTCGGGGGTGTGGGGGCCCGGGCCGCCCTCTCT
CGGAGTTGGCAGGAAGCCAGGTTGCAGGGTGTCCGCTTCCTCAGTTCCAGAGAGGTGGATCG
CATGGTCTCCACGCCCATCGGAGGCCCTCAGCTACGTTCAAGGGTGCACCAAAAAGCATCTTA
ACAGCAAGACTGTGGGCCAGTGCCTGGAGACCACAGCACAGAGGGTCCCAGAACGAGAGGCC
TTGGTCGTCCTCCATGAAGACGTCAGGTTGACCTTTGCCCAACTCAAGGAGGAGGTGGACAA
AGCTGCTTCTGGCCTCCTGAGCATTGGCCTCTGCAAAGGTGACCGGCTGGGCATGTGGGGAC
CTAACTCCTATGCATGGGTGCTCATGTCAGTTGGCCACCGCCCAGGCGGGCATCATTCTGGTG
TCTGTGAACCCAGCCTACCAGGCTATGGAAGTGGAGTATGTCCTCAAGAAGGTGGGCTGCAA
GGCCCTTGTGTTCCCCAAGCAATTCAAGACCCAGCAATACTACAACGTCCTGAAGCAGATCT
GTCCAGAAGTGGAGAATGCCCAGCCAGGGGCCCTTGAAGAGTCAGAGGCTCCCAGATCTGACC
ACAGTCATCTCGGTGGATGCCCCCTTTCGCGGGGACCTGCTCCTGGATGAAGTGGTGGCGGC
TGGCAGCACACGGCAGCATCTGGACCAGCTCCAATACAACCAGCAGTTCCTGTCTGCCATG
ACCCCATCAACATCCAGTTACCTCGGGGACAACAGGCAGCCCCAAGGGGGGCCACCCTCTCC
CACTACAACATTGTCAACAACCTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGAC
ACCAGAGCAGTTGCGGATGATCCTGCCCAACCCCTGTACCATTGCCTGGGTTCCTGGCAG
GCACAATGATGTGTCTGATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGC
AAGAAGGCACTGGAGGCCATCAGCAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGAT
GTTCTGTGGACATTCTGAACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAG
GTGTCAATTGCTGGGTCCCCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAAT
ATGAAGGACCTGGTGGTTGCTTATGGAACCACAGAGAACAGTCCCGTGACATTCGCGCACTT
CCCTGAGGACACTGTGGAGCAGAAGGCAGAAAGCGTGGGCAGAAATTATGCCTCACACGGAGG
CCCGGATCATGAACATGGAGGCAGGGACGCTGGCAAAGCTGAACACGCCCGGGGAGCTGTGC
ATCCGAGGGTACTGCGTCATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGT
GGATCAGGACAAGTGGTATTGGACAGGAGATGTCGCCACAATGAATGAGCAGGGCTTCTGCA
AGATCGTGGGCCGCTCTAAGGATATGATCATCCGGGGTGGTGAGAACATCTACCCCGCAGAG
CTCGAGGACTTCTTTCACACACACCCGAAGGTGCAGGAAGTGCAGGTGGTGGGAGTGAAGGA
CGATCGGATGGGGGAAGAGATTTGTGCCTGCATTGCGCTGAAGGACGGGGAGGAGACCACGG
TGGAGGAGATAAAAGCTTTCTGCAAAGGGAAGATCTCTCACTTCAAGATTCCGAAGTACATC
GTGTTTGTCAAACTACCCCTCACCATTTAGGAAAGATCCAGAAATTCAAACCTCGAGA
GCAGATGGAACGACATCTAAATCTG**TGA**ATAAAGCAGCAGGCCTGTCCTGGCCGGTTGGCTT
GACTCTCTCCTGTCAGAATGCAACCTGGCTTTATGCACCTAGATGTCCCCAGCACCCAGTTC
TGAGCCAGGCACATCAAATGTCAAGGAATTGACTGAACGAACATAAGAGCTCCTGGATGGGTC
CGGGAACCTCGCCTGGGCACAAGGTGCCAAAAGGCAGGCAGCCTGCCCAGGCCCTCCCTCCTG
TCCATCCCCACATTCCCCTGTCTGTCTTGTGATTTGGCATAAAGAGCTTCTGTTTTCTTT
GAAAAAAAAAAAAAAAAA

122/330

FIGURE 122

MAVYVGMLRLGRLCAGSSGVLGARAALSRSWQEARLQGVRFLLSSREVDRMVSTPIGGLSYVQ
GCTKKHLNSKTVGQCLETTAQRVPEREALVVLHEDVRLTFAQLKEEVDKAASGLLSIGLCKG
DRLGMWGPNSYAWVLMQLATAQAGIILVSVNPAYQAMELEYVLKKVGCKALVFPKQFKTQQY
YNVLKQICPEVENAQPGALKSQRLPDLTTVISVDAPLPGTLLLDEVVAAGSTRQHLDQLQYN
QQFLSCHDPINIQFTSGTTGSPKGATLSHYNIVNNSNILGERLKLHEKTPEQLRMILPNPLY
HCLGSVAGTMMCLMYGATLILASPIFNGKKALEAISRRERGTFLYGTPTMFVDILNQPDFSSY
DISTMCGGVIAGSPAPPELIRAIINKINMKDLVVAYGTTENSPVTFAHFPEDTVEQKAESVG
RIMPHTEARIMNMEAGTLAKLNTPGELCIRGYCVMLGYWGEPQKTEEAVDQDKWYWTGDVAT
MNEQGFCCKIVGRSKDMIIRGGENIYPAELEDDFFHTHPKVQEVQVVGKDDRMGEEICACIRL
KDGEETTVEEIKAFCKGKISHFKIPKYIVFVTNYPLTISGKIQKFKLREQMERHLNL

Signal Peptide:

amino acids 1-22

Transmembrane Domains:

amino acids 140-161, 213-229, 312-334

Putative AMP-binding Domain Signature:

amino acids 260-271

N-myristoylation Sites:amino acids 19-24, 22-27, 120-125, 203-208, 268-273, 272-277,
314-319, 318-323, 379-384, 380-385, 409-413**N-glycosylation Site:**

amino acids 282-285

123/330

FIGURE 123

CAACTCCAACATTTTAGGAGAGCGCCTGAAACTGCATGAGAAGACACCAGAGCAGTTGCGGA
TGATCCTGCCCAACCCCCTGTACCATTCGCTGGGTTCCGTGGCAGGCACAATGATGTGTCTG
ATGTACGGTGCCACCCTCATCCTGGCCTCTCCCATCTTCAATGGCAAGAAGGCACTGGAGGC
CATCAGCAGAGAGAGAGAGGCACCTTCCTGTATGGTACCCCCACGATGTTTCGTGGACATTCTGA
ACCAGCCAGACTTCTCCAGTTATGACATCTCGACCATGTGTGGAGGTGTCATTGCTGGGTCC
CCTGCACCTCCAGAGTTGATCCGAGCCATCATCAACAAGATAAATATGAAGGACCTGGTGGT
TGCTTATGGAACCACAGAGAACAGTCCCGTGACATTCGCGCACTTCCCTGAGGACACTGTGG
AGCAGAAGGCAGAAAGCGTGGGCAGAATTATGCCTCACACGGAGGCGCGGATCATGAACATG
GAGGCAGGGACGCTGGCAAAGCTGAACACGCCCCGGGGAGCTGTGCATCCGAGGGTACTGCGT
CATGCTGGGCTACTGGGGTGAGCCTCAGAAGACAGAGGAAGCAGTGGATCAGGACAAGTGGT
ATTGGACAGGAGATGTCGCCAC

124/330

FIGURE 124

GAGCAGGACGGAGCC**ATG**GACCCCGCCAGGAAAGCAGGTGCCCAGGCCATGATCTGGACTGC
AGGCTGGCTGCTGCTGCTGCTGCTTCGCGGAGGAGCGCAGGCCCTGGAGTGCTACAGCTGCG
TGCAGAAAGCAGATGACGGATGCTCCCCGAACAAGATGAAGACAGTGAAGTGCGCGCCGGGC
GTGGACGTCTGCACCGAGGCCGTGGGGGCGGTGGAGACCATCCACGGACAATTCTCGCTGGC
AGTGCGGGGTTGCGGTTTCGGGACTCCCCGGCAAGAATGACCGCGGCCTGGATCTTCACGGGC
TTCTGGCGTTCATCCAGCTGCAGCAATGCGCTCAGGATCGCTGCAACGCCAAGCTCAACCTC
ACCTCGCGGGCGCTCGACCCGGCAGGTAATGAGAGTGCATACCCGCCCAACGGCGTGGAGTG
CTACAGCTGTGTGGGCCTGAGCCGGGAGGCGTGCCAGGGTACATCGCCGCCGGTCTGTGAGCT
GCTACAACGCCAGCGATCATGTCTACAAGGGCTGCTTCGACGGCAACGTCACCTTGACGGCA
GCTAATGTGACTGTGTCCTTGCTGTCCGGGGCTGTGTCCAGGATGAATTCTGCACTCGGGA
TGGAGTAACAGGCCCAGGGTTACGCTCAGTGGCTCCTGTTGCCAGGGGTCCCCTGTAACT
CTGACCTCCGCAACAAGACCTACTTCTCCCCTCGAATCCCACCCCTTGTCCGGCTGCCCCCT
CCAGAGCCCACGACTGTGGCCTCAACCACATCTGTCACCACTTCTACCTCGGCCCCAGTGAG
ACCCACATCCACCACCAAACCCATGCCAGCGCCAACCAGTCAGACTCCGAGACAGGGAGTAG
AACACGAGGCCTCCCGGGATGAGGAGCCCAGGTTGACTGGAGGCGCCGCTGGCCACCAGGAC
CGCAGCAATTCAGGGCAGTATCCTGCAAAAGGGGGGCCCCAGCAGCCCCATAATAAAGGCTG
TGTGGCTCCCACAGCTGGATTGGCAGCCCTTCTGTTGGCCGTGGCTGCTGGTGTCTACTGT**T**
GAGCTTCTCCACCTGGAAATTTCCCTCTCACCTACTTCTCTGGCCCTGGGTACCCCTCTTCT
CATCACTTCCTGTTCCCACTGACTGGGCTGGCCAGCCCCTGTTTTTCCAACATTCCC
CAGTATCCCCAGCTTCTGCTGCGCTGGTTTTCGGCTTTGGGAAATAAAATACCGTTGTATAT
ATTCTGCCAGGGGTGTTCTAGCTTTTTGAGGACAGCTCCTGTATCCTTCTCATCCTTGTCTC
TCCGCTTGTCTCTTGTGATGTTAGGACAGAGTGAGAGAAGTCAGCTGTCACGGGGAAGGTG
AGAGAGAGGATGCTAAGCTTCCTACTCACTTTCTCCTAGCCAGCCTGGACTTTGGAGCGTGG
GGTGGGTGGGACAATGGCTCCCCACTCTAAGCACTGCCTCCCCTACTCCCCGCATCTTTGGG
GAATCGGTTCCCCATATGTCTTCCTTACTAGACTGTGAGCTCCTCGAGGGGGGGCCCGGTAC
CCAATTCGCCCTATAGTGAGTCGTA

125/330

FIGURE 125

MDPARKAGAQAMIWTAGWLLLLLLRGGAQALECYSCVQKADDGCS PNKMKT VKCAPGVDVCT
EAVGAVETIHGQFSLAVRGCGSGLPGKNDRGLDLHGLLAFIQLQOCAQDRCNALNLT SRAL
DPAGNESAYPPNGVECYSCVGLSREACQGTSPPVVSCYNASDHVYKGC FDGNVT LTAA NVTV
SLPVRGCVQDEFCTRDGVTGPGFTLSGSCCQGSRCNSDLRNKTYFSPRI PPLVRLPPPEPTT
VASTTSVTTST SAPVRPTSTTKPMPAPTSQTPRQGEHEASRDEEPRLTGGAAGHQDRS NSG
QYPAKGGPQQPHNKGCVAPTAGLAALLLAVAAGVLL

126/330

FIGURE 126

CGGGACTCGGCGGGTCCTCCTGGGAGTCTCGGAGGGGACCGGCTGTGCAGACGCC**ATG**GAGT
TGGTGCTGGTCTTCCTCTGCAGCCTGCTGGCCCCCATGGTCCTGGCCAGTGCAGCTGAAAAG
GAGAAGGAAATGGACCCTTTTCATTATGATTACCAGACCCTGAGGATTGGGGGACTGGTGTT
CGCTGTGGTCCTCTTCTCGGTTGGGATCCTCCTTATCCTAAGTCGCAGGTGCAAGTGCAGTT
TCAATCAGAAGCCCCGGGCCCCAGGAGATGAGGAAGCCCAGGTGGAGAACCTCATCACCGCC
AATGCAACAGAGCCCCAGAAGCAGAGAACTGAAGTGAGCCATCAGGTGGAAGCCTCTGGAA
CCTGAGGCGGCTGCTTGAACCTTTGGATGCAAATGTGCATGCT**TAA**GAAAACCGGCCACTTC
AGCAACAGCCCTTTCCCCAGGAGAAGCCAAGAACTTGTGTGTCCCCCACCTATCCCCTCTA
ACACCATTCTCCACCTGATGATGCAACTAACACTTGCCTCCCCACTGCAGCCTGCGGTCTT
GCCACCTCCCGTGATGTGTGTGTGTGTGTGTGTGTGTGACTGTGTGTGTTTGCTAACTGTG
GTCTTTGTGGCTACTTGTGTGTGGATGGTATTGTGTTTGTTAGTGAAGTGTGGACTCGCTTT
CCCAGGCAGGGGCTGAGCCACATGGCCATCTGCTCCTCCCTGCCCCCGTGGCCCTCCATCAC
CTTCTGCTCCTAGGAGGCTGCTTGTTGCCCGAGACCAGCCCCCTCCCCTGATTTAGGGATGC
GTAGGGTAAGAGCACGGGCAGTGGTCTTCAGTCGTCTTGGGACCTGGGAAGGTTTGCAGCAC
TTTGTTCATCATTCTTCATGGACTCCTTTCACTCCTTTAAACAAAACCTTGCTTCCTTATCCC
ACCTGATCCCAGTCTGAAGGTCTCTTAGCAACTGGAGATACAAAGCAAGGAGCTGGTGAGCC
CAGCGTTGACGTCAGGCAGGCTATGCCCTTCCGTGGTTAATTTCTTCCCAGGGGCTTCCACG
AGGAGTCCCCATCTGCCCCGCCCCCTTCACAGAGCGCCCGGGGATTCCAGGCCCAGGGCTTCT
ACTCTGCCCCCTGGGGAATGTGTCCCCTGCATATCTTCTCAGCAATAACTCCATGGGCTCTGG
GACCCTACCCCTTCCAACCTTCCCTGCTTCTGAGACTTCAATCTACAGCCCAGCTCATCCAG
ATGCAGACTACAGTCCCTGCAATTGGGTCTCTGGCAGGCAATAGTTGAAGGACTCCTGTTCC
GTTGGGGCCAGCACACCGGGATGGATGGAGGGAGAGCAGAGGCCTTTGCTTCTCTGCCTACG
TCCCCTTAGATGGGCAGCAGAGGCAACTCCCGCATCCTTTGCTCTGCCTGTCGGTGGTCAGA
GCGGTGAGCGAGGTGGGTGGAGACTCAGCAGGCTCCGTGCAGCCCTTGGGAACAGTGAGAG
GTTGAAGGTCATAACGAGAGTGGGAACTCAACCCAGATCCCGCCCCTCCTGTCTCTGTGTT
CCCGCGGAAACCAACCAAACCGTGCGCTGTGACCCATTGCTGTTCTCTGTATCGTGATCTAT
CCTCAACAACAACAGAAAAAAGGAATAAAATATCCTTTGTTTCCT

127/330

FIGURE 127

MELVLVFLCSLLAPMVLASAAEKEKEMDPFHYDYQTLRIGGLVFAVVLFSVGILLILSRCK
CSFNQKPRAPGDEEAQVENLITANATEPQKQRTEVQPSGGSLWNLRRLLLEPLDANVDA

128/330

FIGURE 128

AAACTTGACGCC**ATG**AAGATCCCGGTCCTTCCTGCCGTGGTGCTCCTCTCCCTCCTGGTGCT
CCACTCTGCCCAGGGAGCCACCCTGGGTGGTCCTGAGGAAGAAAGCACCATTGAGAATTATG
CGTCACGACCCGAGGCCTTTAACACCCCGTTCCTGAACATCGACAAATTGCGATCTGCGTTT
AAGGCTGATGAGTTCCTGAACTGGCACGCCCTCTTTGAGTCTATCAAAAGGAACTTCCTTT
CCTCAACTGGGATGCCTTTCCTAAGCTGAAAGGACTGAGGAGCGCAACTCCTGATGCCCAG**T**
GACCATGACCTCCACTGGAAGAGGGGGCTAGCGTGAGCGCTGATTCTCAACCTACCATAACT
CTTTCCTGCCTCAGGAACTCCAATAAAACATTTTCCATCCAAA

129/330

FIGURE 129

MKIPVLPVVLLSLLVLHSAQGATLGGPEEESTIENYASRPEAFNTPFLNIDKLRSFAKDE
FLNWHALFESIKRKLPFLNWDAFPKLGKGLRSATPDAQ

130/330

FIGURE 130

CAGTTCTGAAATCAATGGAGTTAATTTAGGGAATACAAACCAGCC**ATG**GGGGTGGAGATTGC
CTTTGCCTCAGTGATTCTCACCTGCCTCTCCCTTCTGGCAGCAGGAGTCTCCCAGGTTGTTC
TTCTCCAGCCAGTTCCAACCTCAGGAGACAGGTCCCAAGGCCATGGGAGATCTCTCCTGTGGC
TTTGCCGGCCACTCA**TGA**GAGTGTTTTTGTGTAAAGTATTTTTTTAGAATACTGTTGACTTCT
TCATGATTTAATAACCATCCTTTGCGAAGTTTTATGAGGCTTTAGGGGAATGTCAACCCTCA
AATTTTTGTTATACTAGATGGCTTCCATTTACCCACCACTATTTTAAGGTCCCTTTATTTTT
AGGTTCAAGGTTCAATTTGACTTGAGAAAGTGCCCTTCTGCAGCTTCATTGATTTTGTTTATC
TTCATTAATTGTAACGATTAAAAAGAATAAGAGCACGCAGACCTCTAGGAGAATATTT
TATCCCTGGGTGCCCCTGACACATTTATGTAGTGATCCCACAAATGTGATTGTTAATTTAAA
TGTTATTCTAATATTAGTACATTCAGTTGTGATGTAATATGAATAACCAGAATCTATTTCTT
AAAAGTTTTGAGTATATTTTTCAACTAGATATTTGTATAGAAAGACTGAATAGTGATG

131/330

FIGURE 131

MGVEIAFASVILTCLSLAAGVSQVLLQPVPTQETGPKAMGDLSCGFAGHS

132/330

FIGURE 132

GGGGAATCTGCAGTAGGTCTGCCGGCG**ATG**GAGTGGTGGGCTAGCTCGCCGCTTCGGCTCTG
GCTGCTGTTGTTCCCTCCTGCCCTCAGCGCAGGGCCGCCAGAAGGAGTCAGGTTCAAAATGGA
AAGTATTTATTGACCAAATTAACAGGTCTTTGGAGAATTACGAACCATGTTCAAGTCAAAAC
TGCAGCTGCTACCATGGTGTGCATAGAAGAGGATCTAACTCCTTTCCGAGGAGGCATCTCCAG
GAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATCACTAAGAACA
GACTGTACCGGGAAAATGACTGCATGTTCCCCTCAAGGTGTAGTGGTGTGAGCACTTTATT
TTGGAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGATCAATGTACGAGATTATCCTCA
GGTTCCTAAATGGATGGAGCCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACC
ATGATATCATGTATCCTGCTTGGACATTTTGGGAAGGGGGACCTGCTGTTTGGCCAATTTAT
CCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTGGTAAGGTCAGCAGCACAGTG
GCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCAGAACGAG
ATCCTCTCATTCTTCTGTCTCGGAAAAACCCAAAACCTTGTTGATGCAGAATACACCAAAAAC
CAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGT
GGATCACTGCAAATACAAGTATCTGTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTA
AACACCTCTTCCTGTGTGGCTCACTTGTTTTCATGTTGGTGATGAGTGGCTAGAATTCTTC
TATCCACAGCTGAAGCCATGGGTTCATATATCCCAGTCAAAACAGATCTCTCCAATGTCCA
AGAGCTGTTACAATTTGTAAAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGGGGAA
GCCAGTTTATTAGGAACCATTTGCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTG
AGTGAATACTCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTAT
TCCCAAAATGTTGAAAACCTGAACTA**TAG**TAGTCATCATAGGACCATAGTCCTCTTTGTGGCA
ACAGATCTCAGATATCCTACGGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATA
TCTGCTATCAAGCCAAATACCTGGTTTTCTTATCATGCTGCACCCAGAGCAACTCTTGAGA
AAGATTTAAAATGTGTCTAATACACTGATATGAAGCAGTTCAACTTTTTGGATGAATAAGGA
CCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTTCATTTTCTTAAGACCAATC
ACAGCTTGTGCCTCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCA
TGTGATGATGCCCTTTGTCCCATTATTTGGAGCAGAAAATTCGTCATTTGGAAGTAGTACAA
CTCATTGCTGGAATTGTGAAATTATTCAAGGCGTGATCTCTGTCACTTTATTTTAATGTAGG
AAACCTATGGGGTTTATGAAAAATACTTGGGGATCATTCTCTGAATGGTCTAAGGAAGCGG
TAGCCATGCCATGCAATGATGTAGGAGTTCTCTTTTGTAAAACCATAAACTCTGTTACTCAG
GAGGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATTGCAATTGGATT
TCAGGTTCCCTTTTTGTGCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA

133/330

FIGURE 133

MEWWASSPLRLWLLLFLLP
SAQGRQKESGSKWKVFIDQ
INRSLENYEPCSSQNCSCY
HGVIE
EDLTPFRGGISRKMMAEV
VRRKLGTHYQITKNRLYRE
NDCMFPSRCSGVEHFILEV
IGRLPD
MEMVINVRDYPQVPKWME
PAIPVFSFSKTSEYHDIMY
PAWTFWEGGPAVWPIYPT
GLGRWDL
FREDLVRSAQWPWKKNST
AYFRGSRTSPERDPLILLS
SRKNPKLVDAEYTKNQAWK
SMKDT
LGKPAAKDVHLVDHCKYK
YLFNFRGVAASFRFKHLF
LCGSLVFHVGDEWLEFFYP
QLKPWVH
YIPVKTDL
SNVQELLQ
FVKANDD
VAQEIAE
RGSQFIR
NHLQMD
DITCYW
ENLLSE
YSKFL
SY
NVTRRK
GYDQII
PKMLK
TEL

134/330

FIGURE 134

CACCCCTCCATTTCTCGCC**ATG**GGCCCTGCACTGCTCCTGATCCCTGCTGCCCTCGCCTCTT
TCATCCTGGCCTTTGGCACCGGAGTGGAGTTCGTGCGCTTTACCTCCCTTCGGCCACTTCTT
GGAGGGATCCCGGAGTCTGGTGGTCCGGATGCCCCGCCAGGGATGGCTGGCTGCCCTGCAGGA
CCGCAGCATCCTTGCCCCCTGGCATGGGATCTGGGGCTCCTGCTTCTATTTGTTGGGCAGC
ACAGCCTCATGGCAGCTGAAAGAGTGAAGGCATGGACATCCCGGTACTTTGGGGTCCTTCAG
AGGTCACTGTATGTGGCCTGCACTGCCCTGGCCTTGCAGCTGGTGATGCGGTACTGGGAGCC
CATACCCAAAGGCCCTGTGTTGTGGGAGGCTCGGGCTGAGCCATGGGCCACCTGGGTGCCGC
TCCTCTGCTTTGTGCTCCATGTCATCTCCTGGCTCCTCATCTTTAGCATCCTTCTCGTCTTT
GACTATGCTGAGCTCATGGGCCTCAAACAGGTATACTACCATGTGCTGGGGCTGGGCGAGCC
TCTGGCCCTGAAGTCTCCCCGGGCTCTCAGACTCTTCTCCACCTGCGCCACCCAGTGTGTG
TGGAGCTGCTGACAGTGCTGTGGGTGGTGCCTACCCTGGGCACGGACCGTCTCCTCCTTGCT
TTCCTCCTTACCCTCTACCTGGGCCTGGCTCACGGGCTTGATCAGCAAGACCTCCGCTACCT
CCGGGCCCAGCTACAAAGAAACTCCACCTGCTCTCTCGGCCCCAGGATGGGGAGGCAGAGT**G**
AGGAGCTCACTCTGGTTACAAGCCCTGTTCTTCCCTCTCCCACTGAATTCTAAATCCTTAAC
ATCCAGGCCCTGGCTGCTTCATGCCAGAGGCCCAAATCCATGGACTGAAGGAGATGCCCCTT
CTACTACTTGAGACTTTATTCTCTGGGTCCAGCTCCATACCCTAAATTCTGAGTTTCAGCCA
CTGAACTCCAAGGTCCACTTCTCACCAGCAAGGAAGAGTGGGGTATGGAAGTCATCTGTCCC
TTCACTGTTTtagagcatgacactctccccctcaacagcctcctgagaaggaaaggatctgcc
ctgaccactccccctggcactgttacttgcctctgcgccctcaggggtcccccttctgcaccgct
ggcttccactccaagaaggtggaccaggggtctgcaagttcaacgggtcatagctgtccctcca
ggccccaaccttgcctcaccactccccggccctagtctctgcacctccttagggccctgcctct
gggctcagaccccaacctagtcAAGGGGATTCTCCTGCTCTTAACTCGATGACTTGGGGCTC
CCTGCTCTCCCGAGGAAGATGCTCTGCAGGAAAATAAAAGTCAGCCTTTTTCTAAAAAAA

135/330

FIGURE 135

MAPALLLI PAALASFILAFGTGVEFVRFTSLRPLLGGIPESGGPDARQGWLAALQDRSILAP
LAWDLGLLLLFVVGQHSLMAAERVKAWTSRYFGVLQRSLYVACTALALQLVMRYWEPIPKGPV
LWEARAEPWATWVPLLCFVLHVISWLLIFSILLVFDYAELMGLKQVYYHVLGLGEPLALKSP
RALRLFSHLRHPVCVELLTVLWVPTLGTDRLLLAFLLTLYLGLAHGLDQQDLRYLRAQLQR
KLHLLSRPQDGEAE

Signal sequence:

amino acids 1-13

Transmembrane domains:

amino acids 58-76, 99-113, 141-159, 203-222

N-myristoylation sites:

amino acids 37-43, 42-48, 229-235

136/330

FIGURE 136

CCGAGCACAGGAGATTGCCTGCGTTT TAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGA
AGAAATTGCCAAACCATGTCTTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGT
TTTAATTAAGCATGGAATACAGAAAACAACAAAAA ACTTAAGCTTTAATTTTCATCTGGAATT
CCACAGTTTTCTTAGCTCCCTGGACCCGGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTA
TCACGTGGTGCTCTCCGACTACTCACCCCGAGTGTAAGAACCTTCGGCTCGCGTGCTTCTG
AGCTGCTGTGG**ATG**GCCTCGGCTCTCTGGACTGTCCTTCCGAGTAGGATGTCACTGAGATCC
CTCAAATGGAGCCTCCTGCTGCTGCTCACTCCTGAGTTTTCTTTGTGATGTGGTACCTCAGCCT
TCCCCACTACAATGTGATAGAACGCGTGA ACTGGATGTACTTCTATGAGTATGAGCCGATTT
ACAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAACTGCTCTCATCAAAATCCATTT
CTGGTCATTCTGGTGACCTCCCACCCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTAC
TTGGGGTGAAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAG
AGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGT
GACATAATCCGACAAGATTTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTTATGGC
ATTCAGGTGGGTAACTGAGTTTTGCCCCAATGCCAAGTACGTAATGAAGACAGACACTGATG
TTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGAGAAGTTT
TTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATAT
TTCTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTTGGGTTATATAA
TGTCAGAGATTTGGTGCCAAGGATCTATGAAATGATGGGTACGTAAAACCCATCAAGTTT
GAAGATGTTTATGTCGGGATCTGTTTGAATTTATTTAAAGTGAACATTCATATTCCAGAAGA
CACAAATCTTTTCTTTCTATATAGAATCCATTTGGATGTCTGTCAACTGAGACGTGTGATTG
CAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTTTTGGCAGGTCATGCTAAGGAACACC
ACATGCCATTAT**TAA**CTTCACATTCTACAAAAAGCCTAGAAGGACAGGATACCTTGTGGAAA
GTGTTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTCAGTGTGCTGGCTTACACTG
AACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTC
AGGCCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAA
GAAATTAATAGGACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAGGG
TGTTACTGAGTTATAAGCTCACTAGGCTGTAAAAACAAAACAATGTAGAGTTTTATTATTG
AACAAATGTAGTCACTTGAAGGTTTTGTGTATATCTTATGTGGATTACCAATTTAAAAATATA
TGTAGTTCTGTGTCAAAAAACTTCTTCACTGAAGTTATACTGAACAAAATTTTACCTGTTTT
TGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATTATTATTTAAATTA
CTTCAACTTTGTGTTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAGTGAAT
CATTCTTTACATGCAAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCAC
TCCATTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTAAAT
ATTTTACTGTGGTAATATAGAGAAGAATTAAAGCAAGAAAATCTGAAA

137/330

FIGURE 137

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPiYRQD
FHFTLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EK
EDKMLALSLEDEHLLYGDIIRQDFLDTYNNLTLKTIMAFRWVTEFCPNAKYVMKTD TDVFIN
TGNLVKYLLNLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQEYPFKVFPPYCSGLGYIMSRD
LVPRIYEMMGHV KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVCQLRRVIAAHG
FSSKEIITFWQVMLRNTTCHY

138/330

FIGURE 138

CCTCTGTCCACTGCTTTTCGTGAAGACAAGATGAAGTTCACAATTGTCTTTGCTGGACTTCTT
GGAGTCTTTCTAGCTCCTGCCCTAGCTAACTATAATATCAACGTCAATGATGACAACAACAA
TGCTGGAAGTGGGCAGCAGTCAGTGAGTGTCAACAATGAACACAATGTGGCCAATGTTGACA
ATAACAACGGATGGGACTCCTGGAATTCCATCTGGGATTATGGAAATGGCTTTGCTGCAACC
AGACTCTTTCAAAGAAGACATGCATTGTGCACAAAATGAACAAGGAAGTCATGCCCTCCAT
TCAATCCCTTGATGCACTGGTCAAGGAAAAGAAGCTTCAGGGTAAGGGACCAGGAGGACCAC
CTCCCAAGGGCCTGATGTACTCAGTCAACCCAAACAAAGTCGATGACCTGAGCAAGTTCGGA
AAAAACATTGCAAACATGTGTCGTGGGATTCCAACATACATGGCTGAGGAGATGCAAGAGGC
AAGCCTGTTTTTTTTACTCAGGAACGTGCTACACGACCAGTGTACTATGGATTGTGGACATTT
CCTTCTGTGGAGACACGGTGGAGAACTAAACAATTTTTTAAAGCCACTATGGATTTAGTCAT
CTGAATATGCTGTGCAGAAAAAATATGGGCTCCAGTGGTTTTTACCATGTCATTCTGAAATT
TTTCTCTACTAGTTATGTTTGATTTCTTTAAGTTTCAATAAAATCATTTAGCATTGAAAAAAA

139/330

FIGURE 139

MKFTIVFAGLLGVFLAPALANYNINVNDDNNNAGSGQQSVSVNNEHNVANVDNNNGWDSWNS
IWDYGNGFAATRLFQKKTCIVHKMNKEVMPSIQSLDALVKEKKLQKGPGGPPPKGLMYSVN
PNKVDDLKFGKNIANMCRGIPTYMAEEMQEASLFFYSGTCYTTSVLWIVDISFCGDTVEN

Signal Peptide:

amino acids 1-20

N-myristoylation Sites:

amino acids 67-72, 118-123, 163-168

Flavodoxin protein homology:

amino acids 156-174

140/330

FIGURE 140

CATTTCTGAAACTAATCGTGTGAGAATTGACTTTGAAAAGCATTGCTTTTTTACAGAAGTATA
TTAACTTTTTTAGGAGTAATTTCTAGTTTGGATTGTAATATGAAATAATTTAAAAGGGCTTCG
CTCATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTCTTATTGCTTACTGATTTTTT
TGAGTTAAGAGTTGTTATATGCTAGAATATGAGGATGTGAATATAAATAAGAGAAGAAAAA
GAATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAAGCTGGTTTGTTTACATG
CAAGCTTATAGTTGAAATATTTTTTCAGGAATTAC**ATGA**ATGACAGTCTTCGAACCAATGTGT
TTGTTTCGATTTCAACCAGAGACTATAGCATGTGCTTGCATCTACCTTGCAGCTAGAGCACTT
CAGATTCGGTTGCCAACTCGTCCCCATTGGTTTCTTCTTTTTTGGTACTACAGAAGAGGAAAT
CCAGGAAATCTGCATAGAAACACTTAGGCTTTATACCAGAAAAAAGCCAAACTATGAATTAC
TGGAAAAAGAAGTAGAAAAAAGAAAAGTAGCCTTACAAGAAGCCAAATTAAAAGCAAAGGGA
TTGAATCCGGATGGAAGCTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCCAGCCTCCAAGCC
ATCATCACCAAGAGAAGTAAAAGCTGAAGAGAAATCACCAATCTCCATTAATGTGAAGACAG
TCAAAAAAGAACCTGAGGATAGACAACAGGCTTCCAAAAGCCCTTACAATGGTGTAAGAAAA
GACAGCAAGAGAAGTAGAAATAGCAGAAGTGCAAGTCGATCGAGGTCAAGAACACGATCACG
TTCTAGATCACATACTCCAAGAAGACACTATAATAATAGGCGGAGTCGATCTGGAACATACA
GCTCGAGATCAAGAAGCAGGTCCCGCAGTCACAGTGAAAGCCCTCGAAGACATCATAATCAT
GGTTCTCCTCACCTTAAGGCCAAGCATAACCAGAGATGATTTAAAAGTTCAAACAGACATGG
TCATAAAAGGAAAAAATCTCGTTCTCGATCTCAGAGCAAGTCTCGGGATCACTCAGATGCAG
CCAAGAAACACAGGCATGAAAGGGGACATCATAGGGACAGGCGTGAACGATCTCGCTCCTTT
GAGAGGTCCCATAAAAGCAAGCACCATGGTGGCAGTCGCTCAGGACATGGCAGGCACAGGCG
CTGACTTTCTCTTCCTTTGAGCCTGCATCAGTTCTTGGTTTTGCCTATCTACAGTGTGATGT
ATGGACTCAATCAAAAACATTAAACGCAAACCTGATTAGGATTTGATTTCTTGAAACCCTCTA
GGTCTCTAGAACACTGAGGACAGTTTCTTTTGAAAAGAACTATGTTAATTTTTTTGCACATT
AAAATGCCCTAGCAGTATCTAATTAAAAACCATGGTCAGGTTCAATTGTACTTTATTATAGT
TGTGTATTGTTTATTGCTATAAGAACTGGAGCGTGAATTCTGTAAAAATGTATCTTATTTTT
ATACAGATAAAATTGCAGACACTGTTCTATTTAAGTGGTTATTTGTTTAAATGATGGTGAAT
ACTTTCTTAACACTGGTTTGTCTGCATGTGTAAAGATTTTTTACAAGGAAATAAAATACAAAT
CTTGTTTTTTCTAAAAAAGAAAAAAGT

141/330

FIGURE 141

MNDSLRTNVFVRFQPETIACACIYLAARALQIPLPTRPHWFLLFGTTEEEIQEICIETLRLY
TRKKPNYELLEKEVEKRVKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEEK
SPISINVKTVKKEPEDRQQASKSPYNGVRKDSKRSRNSRSASRSRSRTRSRSRSHTPRRHYN
NRRSRSGTYSSRSRSRSRSHSESPRRHHNHGSPHLKAKHTRDDLKSSNRHGHKRKKSRSRSQ
SKSRDHSDAAKKHRHERGHHRRDRRERSRSFERSHKSKHHGGSRSRGHGRHRR

142/330

FIGURE 142

TGGGGATAAAGGAAAAATGGTCAGGTATTAATGGCTTAAAGATTATTGGAAGGGGTTTATCA
TTTTTTGAANNTATTCGGGTCANAATTGNCTTTGAAAAGCATTGCTTTTTACAGAAATATAT
TANCTTTTTAGAGTAATTTCTAGTTTGGATTGTAATATGAAATTATTTAAAAGGGCTTCGCT
CATATATAGGAAAATCGCATATGGTCCTAGTATTAAATTNNTATTGCTTACTGATTTTTTTTG
AGTTAAGAGTTGTTATATGNTAGAATATGAGGATGTGAATATAAATAAGAGAAGAAAAAAGA
ATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAAGTGGTTTGTTTACATGCA
AGCTTATAGTTGAAATATTTTTCAGGAATTACATGAATGACAGTCTTCGAACCAATGTGTTT
GTTCGATTTCAACCAGAGANTATAGCATGTGCTTGCATCTACCTTGCAGNTAGAGCACTTCA
GATTCCGTTGCCAACTNGTCCCCATTGGTTTCTTCTTTTTGGTACTACAGAAGAGGAAATCC
AGGAAATNTGCATAGAAACACTTAGGCTTTATACCAGAAAAAAGCCAACTATGAATTACTG
GAAAAAGAAGTAGAAAAAAGAAAAGTAGCCTTACAAGAAGCCNAATTAAAAGCAAAGGGATT
GAATCCGGATGGAAGTCCAGCCCTTTCAACCCTGGGTGGATTTTCTCC

143/330

FIGURE 143

GGCACGAGGCCCTCGTGCCAAGCTTGGCACGAGGGTGCACCGCGTTCTCGCACGCGTC**ATGGC**
GGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCACCCTCATGCACAGGCTGGCGC
CACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAAACGGCAGTTTGTTCGGATACAAGCACCCG
TCTGAGGAGGAGCTTCGGGGCCCTGGCGGGGAAGCCGAGGCCAGAGGCAGGAAAGAGCGGTG
GGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCCGAGATGCCCCGTTCCAGCTGG
AGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGCTTCTTCCTGGAGTACCAGTGG
TTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTTCACAGAGGCCTACTACTACAT
GCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGTGCCTGCTCACGGTGACCTTCT
CCATCAAGATGTTCCCTGACAGTGACACGGCTGTACTTCAGCGCCGAGGAGGGGGGTGAGCGC
TCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGCCATGCTGGTGCAAGTGGTGCG
GGAGGAGACCCTCGAGCTGGGCCTGGAGCCTGGTCTGGCCAGCATGACCCAGAACTTAGAGC
CACTTCTGAAGAAGCAGGGCTGGGACTGGGCGCTTCCTGTGGCCAAGCTGGCTATCCGCGTG
GGACTGGCAGTGGTGGGCTCTGTGCTGGGTGCCTTCCTCACCTTCCCAGGCCTGCGGCTGGC
CCAGACCCACCGGGACGCACTGACCATGTCTGGAGGACAGACCCATGCTGCAGTTCCTCCTGC
ACACCAGCTTCCTGTCTCCCCTGTTTCATCCTGTGGCTCTGGACAAAGCCCATTGCACGGGAC
TTCCTGCACCAGCCGCCGTTTGGGGAGACGCGTTTCTCCCTGCTGTCCGATTCTGCCTTCGA
CTCTGGGCGCCTCTGGTTGCTGGTGGTGTGCTGTGCCTGCTGCGGCTGGCGGTGACCCGGCCCC
ACCTGCAGGCCTACCTGTGCCTGGCCAAGGCCCGGGTGGAGCAGCTGCGAAGGGAGGCTGGC
CGCATCGAAGCCCGTGAAATCCAGCAGAGGGTGGTCCGAGTCTACTGCTATGTGACCGTGGT
GAGCTTGCAGTACCTGACGCCGCTCATCCTCACCTCAACTGCACACTTCTGCTCAAGACGC
TGGGAGGCTATTCTGGGGCCTGGGCCAGCTCCTCTACTATCCCCCGACCCATCCTCAGCC
AGCGCTGCCCCCATCGGCTCTGGGGAGGACGAAGTCCAGCAGACTGCAGCGCGGATTGCCGG
GGCCCTGGGTGGCCTGCTTACTCCCCTCTTCCTCCGTGGCGTCTGGCCTACCTCATCTGGT
GGACGGCTGCCTGCCAGCTGCTCGCCAGCCTTTTCGGCCTCTACTTCCACCAGCACTTGGCA
GGCTCC**TAG**CTGCCTGCAGACCCTCCTGGGGCCCTGAGGTCTGTTCTGGGGCAGCGGGACA
CTAGCCTGCCCCCTCTGTTTGCGCCCCCGTGTCCCCAGCTGCAAGGTGGGGCCGGACTCCCC
GGCGTTCCTTCACCACAGTGCCTGACCCGCGGCCCCCCTTGGACGCCGAGTTTCTGCCTCA
GAACTGTCTCTCCTGGGCCCAGCAGCATGAGGGTCCCGAGGCCATTGTCTCCGAAGCGTATG
TGCCAGGTTTGAGTGGCGAGGGTGATGCTGGCTGCTCTTCTGAACAAATAAAGGAGCATGCC
GATTTTAA

144/330

FIGURE 144

MAVLGVQLVVTLLTATLMHRLAPHCSFARWLLCNGSLFRYKHPSEEELRALAGKPRPRGRKE
RWANGLSEEKPLSVPRDAPFQLETCPLTTVDALVLRFFLEYQWFVDFAVYSGGVYLFTEAYY
YMLGPAKETNIAVFWCLLTVTFSIKMFLTVTRLYFSAEEGGERSVCLTFAFLFLLLAMLVQV
VREETLELGLEPGLASMTQNLEPLLKKQGWDWALPVAKLAI RVGLAVVGSVLGAFLTFFGLR
LAQTHRDALTMSEDRPMLQFLLHTSFLSPLFILWLWTKPIARDFLHQPPFGETRFSLLSDSA
FDSGRLWLLVVLCLLRLAVTRPHLQAYLCLAKARVEQLRREAGRIEAREIQQRVVRVYCYVT
VVS LQYLTPLILTLNCTLLLKTLGGYSWGLGPAPLLSPDPSSASAAPIGSGEDEVQQTAAARI
AGALGGLLTPLFLRGVLAYLIWWTAACQLLASLFGLYFHQHLA GS

145/330

FIGURE 145

CGTTNGCACGCGTCAATGGCGGTCCTCGGAGTACAGCTGGTGGTGACCCTGCTCACTGCCAC
CCTCATGCACAGGCTGGCGCCACACTGCTCCTTCGCGCGCTGGCTGCTCTGTAAACGGCAGTT
TGTTCCGATACAAGCACCCGTNTTGAGGAGGAGCTTCGGGGCCCTGGCGGGGAAGCCGAGGCC
CAGAGGCAGGAAAGAGCGGTGGGCCAATGGCCTTAGTGAGGAGAAGCCACTGTCTGTGCCCC
GAGATGCCCCGTTCCAGCTGGAGACCTGCCCCCTCACGACCGTGGATGCCCTGGTCCTGCGC
TTCTTCCTGGAGTACCAGTGGTTTGTGGACTTTGCTGTGTACTCGGGCGGCGTGTACCTCTT
CACAGAGGCCTACTACTACATGCTGGGACCAGCCAAGGAGACTAACATTGCTGTGTTCTGGT
GCCTGCTCACAGTGACCTTCTCCATCAAGATGTTCTGACAGTGACACGGCTGTACTTCAGC
GCCGAGGAGGGGGGTGAGCGCTCTGTCTGCCTCACCTTTGCCTTCCTCTTCCTGCTGCTGGC
CATGCTGGTGCAAGCG

146/330

FIGURE 146

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCCGTGATTTATTAACGTGGCTT
AATCTGAAGGTTCTCAGTCAAATTCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGC
TTGGCTGGTTTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGA**ATGA**AGG
CGCTTCTGTTGCTGGTCTTGCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTG
TATTCAGAACTCTGTAAAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTG
TCCAGACGGCTGTGCGAGCCTCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAA
TGACAGACGAGCCTGGCCTAGACAACCCTGCCTACGTGTCCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCA
GTGGACTCTGGCCGGAGCAACCGAACTAGGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAA
AAAAATAAATCGAGCTTTGAGTGTTCTTCGAAGGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGG
GCAGGGAAAATTCTGAAAACACCACTGCCCCCTGAAGTCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAA
ATTACCAGCATCAAGATCAATCGAGTAGATCCCAGTGAAAGCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAAC
CCCCTGGTCCATATCATTATCCAACACATTTATCGTGATGGGGTGATCGCCAGAGACGGCCGGCTACTGCCAG
GAGACATCATTTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACTACGCTGTGCGTCTCCTGCGG
CAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCGCGAGCAGGAACAATGGACAGGCCCC
GGATGCCTACAGACCCCCGAGATGACAGCTTTTCATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGAA
TAAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTTCATCTTCAATGTGCTGGATGGCGGTGTGGCATATCGA
CATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCAGCCCAGAAAG
TGCGGCTCATCTGATTACAGGCCAGTGAAAGACGTGTTACCTCGTCGTGTCCCGCCAGGTTCCGGCAGCGGAGCC
CTGACATCTTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACACT
CCCAAGCCCCCTCCATCCTACAATTACTTGTTCATGAGAAGGTGGTAAATATCCAAAAAGACCCCGGTGAATCTCT
CGGCATGACCGTTCGACGGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCATCAGTGTTGAGCCCG
GAGGAGTCATAAGCAGAGATGGAAGAATAAAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCGAAGTACA
GAGGTACAGCCGGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGT
CAAAGAGTATGAGCCCCAGGAAGACTGCAGCAGCCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCA
GTGACTGGTCCCCATCCTGGGTTCATGTGGCTGGAATTACCACGGTGCTTGATAACTGTAAAGATATTGTATTA
CGAAGAAACACAGCTGGAAGTCTGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAAACAAACCTTT
TTTCATCAAATCCATTGTTGAAGGAACACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTG
CTGTCAATGGTAGAAGTACATCAGGAATGATACATGCTTGCTTGGCAAGACTGCTGAAAGAAGTTAAAGGAAGA
ATTACTCTAACTATTGTTTCTTGGCCTGGCACTTTTTTAT**AGA**ATCAATGATGGGTGAGAGGAAAACAGAAAA
TCACAAATAGGCTAAGAAGTTGAAACACTATATTTATCTTGTGATTTTTATATTTAAAGAAAGAATACATTGT
AAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAATTA
AAACTACTAGTTTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTACAACATTGTTTATATTTTTTCTATTCAAT
AAAAAGCCCTAAAACAATAAATGATTGATTTGTATACCCCACTGAATTCAGCTGATTTAAATTTAAATTT
GGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTTTAATTTACAGCTAAAATATTTTTTAAATGCA
TTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATATTTTTTCAGAAGTTAAA

147/330

FIGURE 147

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPDGCASLTAT
APSPEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRS
FKKINRALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDP
SELSIRLVGGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLL
RQPCQVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGV
FIFNVLDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQVRQRS
PDIFQEAGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHRE
WDLPIYVISVEPGGVISR DGRIKTGDILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEV
KEYEPQEDCSSPAALDSNHNMAPPSDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIV
GGYEEYNGNKPFFIKSIVEGTPAYNDGRIRCGDILLAVNGRSTSGMIHACLARLLKELKGRI
TLTIVSWPGTFL

148/330

FIGURE 148

CCAAAGTGATCATTTGAAAAAGAGATATCCACATCTTCAAGCCCATATAAAGGATAGAAGCT
GCACAGGGCAGCTTTACTTACTCCAGCACCTTCCTCTCCCAGGCAA**ATG**GTGCTGACCATCT
TTGGGATACAATCTCATGGATACGAGGTTTTTAAACATCATCAGCCCAAGCAACAATGGTGGC
AATG TTCAGGAGACAGTGACAATTGATAATGAAAAAATAACCGCCATCGTTAACATCCATGC
AGGATCATGCTCTTCTACCACAATTTTTGACTATAAACATGGCTACATTGCATCCAGGGTGC
TCTCCCGAAGAGCCTGCTTTATCCTGAAGATGGACCATCAGAACATCCCTCCTCTGAACAAT
CTCCAATGGTACATCTATGAGAAACAGGCTCTGGACAACATGTTCTCCAACAAATACACCTG
GGTCAAGTACAACCCTCTGGAGTCTCTGATCAAAGACGTGGATTGGTTCCTGCTTGGGTCAC
CCATTGAGAACTCTGCAAACATATCCCTTTGTATAAGGGGGAAGTGGTTGAAAACACACAT
AATGTCGGTGCTGGAGGCTGTGCAAAGGCTGGGCTCCTGGGCATCTTGGGAATTTCAATCTG
TGCAGACATTCATGTT**TAG**GATGATTAGCCCTCTTGTTTTATCTTTTCAAAGAAATACATCC
TTGGTTTACACTCAAAGTCAAATTAAATTCTTTCCCAATGCCCCAACTAATTTTGAGATTC
AGTCAGAAAATATAAATGCTGTATTTATA

149/330

FIGURE 149

MKILVAFLVVLTIQSHGYEVFNIIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTT
IFDYKHGYIASRVLSRRACFILKMDHQNIPLNNLQWYIYEKQALDNMFSNKYTWVKYNPLE
SLIKDVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

150/330

FIGURE 150

GGCACGAGCCAGGAACTAGGAGGTTCTCACTGCCCCGAGCAGAGGCCCTACACCCACCGAGGC
ATGGGGCTCCCTGGGCTGTTCTGCTTGCCGTGCTGGCTGCCAGCAGCTTCTCCAAGGCACG
GGAGGAAGAAATTACCCCTGTGGTCTCCATTGCCTACAAAGTCCTGGAAGTTTTCCCCAAAG
GCCGCTGGGTGCTCATAACCTGCTGTGCACCCCAGCCACCACCGCCCATCACCTATTCCCTC
TGTGGAACCAAGAACATCAAGGTGGCCAAGAAGGTGGTGAAGACCCACGAGCCGGCCTCCTT
CAACCTCAACGTCACACTCAAGTCCAGTCCAGACCTGCTCACCTACTTCTGCCGGGCGTCCT
CCACCTCAGGTGCCCATGTGGACAGTGCCAGGCTACAGATGCACTGGGAGCTGTGGTCCAAG
CCAGTGTCTGAGCTGCGGGCCAACCTTCACTCTGCAGGACAGAGGGGCAGGCCCCAGGGTGGA
GATGATCTGCCAGGCGTCCTCGGGCAGCCCACCTATCACCAACAGCCTGATCGGGAAGGATG
GGCAGGTCCACCTGCAGCAGAGACCATGCCACAGGCAGCCTGCCAACTTCTCCTTCCTGCCG
AGCCAGACATCGGACTGGTTCTGGTGCCAGGCTGCAAACAACGCCAATGTCCAGCACAGCGC
CCTCACAGTGGTGCCCCCAGGTGGTGACCAGAAGATGGAGGACTGGCAGGGTCCCCTGGAGA
GCCCCATCCTTGCCCTTGCCGCTCTACAGGAGCACCCGCCGTCTGAGTGAAGAGGAGTTTGGG
GGGTTCAGGATAGGGAATGGGGAGGTCAGAGGACGCAAAGCAGCAGCCATG**TAGA**ATGAACC
GTCCAGAGAGCCAAGCACGGCAGAGGACTGCAGGCCATCAGCGTGCACTGTTTCGTATTTGGA
GTTTCATGCAAAATGAGTGTGTTTTAGCTGCTCTTGCCACAAAAAAAAAAAAAAAAAAAAA

151/330

FIGURE 151

MGLPGLFCLAVLAASSFSKAREEEITPVVSIAYKVLEVFPKGRWVLITCCAPQPPPPITYSL
CGTKNIKVAKKVVKTHEPASFNLNVTLKSSPDLLTYFCRASSTSGAHVDSARLQMHWELWSK
PVSELNANFTLQDRGAGPRVEMICQASSGSPITNSLIGKDGQVHLQQRPCHRQPANFSFLP
SQTSDWFWCQAANNANVQHSALTVVPPGGDQKMEDWQGPLESPIALALPLYRSTRRLSEEEFG
GFRIGNGEVRGRKAAAM

Signal Peptide:

amino acids 1-18

N-glycosylation Sites:

amino acids 86-89, 132-135, 181-184

152/330

FIGURE 152

GGTCCTTA**ATG**GCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTG
CTGTCCGGCTGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGT
CATCCCTAAGTTCAGACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGA
CTTTTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAAA
CTAAATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACT
TACAGAGCAACTGCGTGACATTGAGCTGGAGAATTACACACCCAAGGAACCCCTCACCCCTGC
AGGCAAGGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGT
TTCGATGGGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCC
TGGAGCCAGAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATT
ACTTCTCAATGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACC
CTGGAGCCAAGTGCAGGAGCACCACTCGCCATGTCCTCAGGCACAACCCAACCTCAGGGGCCAC
AGCCACCACCCTCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCA
TCTGAGGAGAGTCCTTTAGAGTGACAGGTTAAAGCTGATACCAAAGGCTCCTGTGAGCAG
GTCTTGATCAAACCTCGCCCTTCTGTCTGGCCAGCTGCCCACGACCTACGGTGTATGTCCAGT
GGCCTCCAGCAGATCATGATGACATCATGGACCCAATAGCTCATTCCTGACCTTGATTCCCTT
TTGCCAACAATTTTACCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACC
TGATGGAATTCCTGCACTTAAAGTTCTGGCTGACTAAACAAGATATATCATTTTCTTTCTTC
TCTTTTGTGGAAAATCAAGTACTTCTTTGAATGATGATCTCTTTCTTGCAAATGATATT
GTCAGTAAAATAATCACGTTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAG
AATTTTTAAATTATTTAATAAGAAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTAT
TGTTCTGTACTGATATTTAATAAAGAGTTCTATTTCCCAAAAAAAAAAAAAAAAAAAAA

153/330

FIGURE 153

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTF
HYDCGNKTVTPVSPLGKKLNVTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQAR
MSCEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFS
MGDCIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLLCCLLIILPCFILPGI

Important features:**Signal peptide:**

amino acids 1-25

Transmembrane domain:

amino acids 224-246

N-glycosylation site.

amino acids 68-72, 82-86

N-myristoylation site.

amino acids 200-206, 210-216

Amidation site.

amino acids 77-81

154/330

FIGURE 154

GGGAAAGCCATTTGAAAACCCATCTATACAACTATATATTTTCATTTCTGCTGCTAGCTG
CCTTGGGCCTCACAATTTTCATTCTGTTTTCTGACTTTCAAGTTATATACCGTGGA**ATG**GAG
TTGATCCCAACCATAACATCGTGGAGGGTTTTAATTTTGGTGGTAGCCCTCACCCAATTCTG
GTGTGGCTTTCTTTGCAGAGGATTCCACCTTCAAAATCATGAACTCTGGCTGTTGATCAAAA
GAGAATTTGGATTCTACTCTAAAAGTCAATATAGGACTTGGCAAAGAAGCTAGCAGAAGAC
TCAACCTGGCCTCCCATAAACAGGACAGATTATTCAGGTGATGGCAAAAATGGATTCTACAT
CAACGGAGGCTATGAAAGCCATGAACAGATTCCAAAAAGAAAACCTCAAATTGGGAGGCCAAC
CCACAGAACAGCATTTCTGGGCCAGGCTG**TAA**TCAGAATTGTCGTCGTACATGCTCAACAGC
ATTGCTTTTTTCCCCAAAATTAACACATTGTGGAGAAGTGATGATACTCTCCCCTTACCTTT
CCTCTCTCCATTCAAGCATTCAAAGTATATTTTCAATGAATTAAACCTTGCAGCAAGGGACC
TTAGATAGGCTTATTCTGACTGTATGCTTTACCAATGAGAGAAAAAAATGCATTTCTGTAT
CATCCTTTTCAATAAACTGTATTCATTTTGAAAAAAAAAAAAAAAAAAAAA

155/330

FIGURE 155

MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHFWLLIKREFGFYSKSQYRTWQKKLA
EDSTWPPINRTDYSGDGKNGFYINGGYESHEQIPKRKLKLGGQPTEQHFWARL

156/330

FIGURE 156

GTTCTCCTTTCCGAGCCAAAATCCCAGGCGATGGTGAATTATGAACGTGCCACACC**ATGAAG**
CTCTTGTTGGCAGGTAAGTGTGCACCACCACACCTGGAATGCCATCCTGCTCCCGTTTCGTCTA
CCTCACGGCGCAAGTGTGGATTCTGTGTGCAGCCATCGCTGCTGCCGCCTCAGCCGGGCCCC
AGAACTGCCCCCTCCGTTTGCTCGTGCAGTAACCAAGTTCAGCAAGGTGGTGTGCACGCGCCGG
GGCCTCTCCGAGGTCCCGCAGGGTATTCCCTCGAACACCCGGTACCTCAACCTCATGGAGAA
CAACATCCAGATGATCCAGGCCGACACCTTCCGCCACCTCCACCACCTGGAGGTCTGCAGT
TGGGCAGGAAGTCCATCCGGCAGATTGAGGTGGGGGCCTTCAACGGCCTGGCCAGCCTCAAC
ACCCTGGAGCTGTTTCGACAACTGGCTGACAGTCATCCCTAGCGGGGCCTTTGAATACCTGTC
CAAGCTGCGGGAGCTCTGGCTTCGCAACAACCCCATCGAAAGCATCCCCTCTTACGCCTTCA
ACCGGGTGCCCTCCCTCATGCGCCTGGACTTGGGGGAGCTCAAGAAGCTGGAGTATATCTCT
GAGGGAGCTTTTGAGGGGCTGTTCAACCTCAAGTATCTGAACTTGGGCATGTGCAACATTAA
AGACATGCCCAATCTCACCCCCCTGGTGGGGCTGGAGGAGCTGGAGATGTCAGGGAACCACT
TCCCTGAGATCAGGCCTGGCTCCTTCCATGGCCTGAGCTCCCTCAAGAAGCTCTGGGTCTAG
AACTCACAGGTCAGCCTGATTGAGCGGAATGCTTTTGACGGGCTGGCTTCACTTGTGGAAGT
CAACTTGGCCCACAATAACCTCTCTTCTTTGCCCCATGACCTCTTTACCCCGCTGAGGTACC
TGGTGGAGTTGCATCTACACCACAACCTTGGAACTGTGATTGTGACATTCTGTGGCTAGCC
TGGTGGCTTCGAGAGTATATACCCACCAATTCCACCTGCTGTGGCCGCTGTCATGCTCCCAT
GCACATGCGAGGCCGCTACCTCGTGGAGGTGGACCAGGCCTCCTTCCAGTGCTCTGCCCCCT
TCATCATGGACGCACCTCGAGACCTCAACATTTCTGAGGGTCGGATGGCAGAACTTAAGTGT
CGGACTCCCCCTATGTCTCCTCCGTGAAGTGGTTGCTGCCCAATGGGACAGTGCTCAGCCACGC
CTCCCGCCACCCAAGGATCTCTGTCTCAACGACGGCACCTTGAACCTTTTCCACGTGCTGC
TTTCAGACACTGGGGTGTACACATGCATGGTGACCAATGTTGCAGGCAACTCCAACGCCTCG
GCCTACCTCAATGTGAGCACGGCTGAGCTTAACACCTCCAACCTACAGCTTCTTACCACAGT
AACAGTGGAGACCACGGAGATCTCGCCTGAGGACACAACGCGAAAGTACAAGCCTGTTCTTA
CCACGTCCACTGGTTACCAGCCGGCATATACACCTCTACCACGGTGCTCATTACAGACTACC
CGTGTGCCCAAGCAGGTGGCAGTACCCGCGACAGACACCACTGACAAGATGCAGACCAGCCT
GGATGAAGTCATGAAGACCACCAAGATCATCATTGGCTGCTTTGTGGCAGTGACTCTGCTAG
CTGCCGCCATGTTGATTGTCTTCTATAAACTTCGTAAGCGGCACCAGCAGCGGAGTACAGTC
ACAGCCGCCCCGACTGTTGAGATAATCCAGGTGGACGAAGACATCCCAGCAGCAACATCCGC
AGCAGCAACAGCAGCTCCGTCCGGTGTATCAGGTGAGGGGGCAGTAGTGCTGCCCACAATTC
ATGACCATATTAACATAACACCTACAAACCAGCACATGGGGCCCACTGGACAGAAAACAGC
CTGGGGAACTCTCTGCACCCACAGTCACCACTATCTCTGAACCTTATATAATTACAGACCCA
TACCAAGGACAAGGTACAGGAACTCAAATA**TGA**CTCCCCTCCCCCAAAAACTTATAAAAT
GCAATAGAATGCACACAAAGACAGCAACTTTTGTACAGAGTGGGGAGAGACTTTTTCTTGTA
TATGCTTATATATTAAGTCTATGGGCTGGTTAAAAAAAACAGATTATATTAAATTTAAAGA
CAAAAAGTCAAAACA

157/330

FIGURE 157

MKLLWQVTVHHHTWNAILLPFVYLTAQVWILCAAIAAAASAGPQNCPSVCSCSNQFSKVVCT
RRGLSEVPQGIPSNTRYLNLMENNIQMIQADTFRHLHHLEVLQLGRNSIRQIEVGAFNGLAS
LNTLELFDNWLTVIPSGAFEYLSKLRRLWLRNNPIESIPSYAFNRVP SLMRLDLGELKKLEY
ISEGAFEGLEFNLKYLNLGMCNIKDMPNLTPLVGLEELEM SGNHFPEIRPGSFHGLSSLKKLW
VMNSQVSLIERN AFDGLASLVELNLAHNNLSSLP HDLFTPLRYLVELHLHHPWNCD CDILW
LAWWLREYIPTNSTCCGRCHAPMHMRGRYLVEVDQASFQCSAPFIMDAPRDLNISEGRMAEL
KCRTPPMSSSVKWLLPNGTVLSHASRHPRISVLNDGT LNF SHVLLSDTGVYTCMVTNVAGNSN
ASAYLNVSTAELNTSNYSFFT TTVTVETTEISPEDTTRKYKPVPTTSTGYQPAYTTSTTVLIQ
TTRVPKQVAVPATDTTDKMQTSLDEVMKTTKIIIGCFVAVTLLAAAMLIVFYKLRKRHQQRS
TVTAARTVEIIQVDEDIPAATSAAATAAPSGVSGEGAVVLPTIHDHINYNTYKPAHGAHWTE
NSLGNLHPTVTTISEPYIIQTHTKDKVQETQI

FIGURE 158

[illegible]

159/330

FIGURE 159

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVV
GYTIPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDEFYVKGIFYCAECRAGWYGGD
CMRCGQVLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRD
GDNRDGQIIKRVCGNERPAPIQSIGSSLHVLHFSDGSKNFDGFHAIYEEITACSSSPCFHDG
TCVLDKAGSYKCACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTV
VSFFCNNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPMQVQSRETPLH
QLYSAAFSKQKLQSAPTKKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGK
WSGRAPSCIPICGKIENITAPKTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNE
RTVVVAAHCVTDLGKVTMIKTADLKVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLD
ADIAILKLLDKARISTRVQPICLAASRDLSTSFQESHITVAGWNVLADVRS PGFKNDTLRSG
VVSVDSSLCEEQHEDHGIPVSVTDNMFCASWEPTAPSDICTAETGGIAAVSFPGRASPEPR
WHLMGLVSWSYDKTCSHRLSTAFTKVLPFKDWIERNMK

160/330

FIGURE 160

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGA
AGCTTTCTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAAC**ATG**GGC
TTCAACCTGACTTTCCACCTTTCCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCT
GACAGTGGTTGGGTGGGCCACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCCTAAAG
CAAAGGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAAAACTCTGACT
AATGAAGCATCCACGAAGAAGGTAGAACTTGACAACTGTCCTTCTGTGTCTCCTTACCTCAG
AGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATC
CCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCGCCATC
CTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAACATCTGCATCCCTT
CCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGGTAAAAAGT
TTAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGGAC
TGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGA
GGAGCATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTG
GATATTTTGGGGGTGTTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCT
AACAACTACTGGGGATGGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAG
AATGAAAATTTCCCGGCCCTGCCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAG
ACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTACAGAGTCTGG
AGAACAGATGGGTTGAGTAGTTGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATA
TATCAACATCACAGTGGATTTCTGGTTTGGTGCA**TGA**CCCTGGATCTTTTGGTGATGTTTGG
AAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTA
AGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCCTTTTTGTATTTTCTTAGCAGAGCT
CCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTCTTAGTCATTTTGATCATG
AGGGTTAAATATTGTAATATGGTACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGAT
AAAATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATAT
TATGGGATAAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCT
CGTCCAAGGTAGAAAGGTACGAAGATACAATACTGTTATTCATTTATCCTGTACAATCATCT
GTGAAGTGGTGGTGTGAGGTGAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCA
GGACACAGTGAACCTGGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTGCGCTGCAAAGGCAG
CAGTAGCTGAGCTGGTTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCCAGGTATGCCT
TCCAGTGATGCCCACCAGAGAATACATTCTCTATTAGTTTTTAAAGAGTTTTTGTAAAATGA
TTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACATATTAATAATAATAATA
TGTCTATCAAATACCTCTGTAGTAAAATGTGAAAAAGCAAAA

161/330

FIGURE 161

MGFNLT FHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKT
LTNEASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRV
AILVPHRNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKGKFNRAKLLNVGYLEALKEEN
WDCFIFHDVDLVPENDFNLYKCEEHPKHLVGRNSTGYRLRYSGYFGGVTALSREQFFKVNG
FSNNYWGWGGEDDDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSR
VWRTDGLSSCSYKLVSVEHNPLYINITVDFWFGA

Important features:**Signal peptide:**

amino acids 1-27

N-glycosylation sites:

amino acids 4-7, 220-223 and 335-338

Xylose isomerase proteins:

amino acids 191-201

162/330

FIGURE 162

CGTGGGCCGGGGTTCGCGCAGCGGGCTGTGGGCGCGCCCGGAGGAGCGACCGCCGCAGTTCTC
GAGCTCCAGCTGCATTCCCTCCGCGTCCGCCCCACGCTTCTCCCGCTCCGGGGCCCCGCA**ATG**
GCCCAGGCAGTGTGGTTCGCGCCTCGGCCGCATCCTCTGGCTTGCCTGCCTCCTGCCCTGGGC
CCCGGCAGGGGTGGCCGCAGGCCTGTATGAACTCAATCTCACCACCGATAGCCCTGCCACCA
CGGGAGCGGTGGTGACCATCTCGGCCAGCCTGGTGGCCAAGGACAACGGCAGCCTGGCCCTG
CCCGCTGACGCCCACCTCTACCGCTTCCACTGGATCCACACCCCGCTGGTGTCTTACTGGCAA
GATGGAGAAGGGTCTCAGCTCCACCATCCGTGTGGTTCGGCCACGTGCCCGGGGAATTCCCGG
TCTCTGTCTGGGTCACTGCCGCTGACTGCTGGATGTGCCAGCCTGTGGCCAGGGGCTTTGTG
GTCCTCCCCATCACAGAGTTCCCTCGTGGGGGACCTTGTTGTACCCAGAACAACCTTCCCTACC
CTGGCCCAGCTCCTATCTCACTAAGACCGTCCTGAAAGTCTCCTTCCCTCCACGACCCGA
GCAACTTCCCTCAAGACCGCCTTGTTTCTCTACAGCTGGGACTTCGGGGACGGGACCCAGATG
GTGACTGAAGACTCCGTGGTCTATTATAACTATTCCATCATCGGGACCTTCACCGTGAAGCT
CAAAGTGGTGGCGGAGTGGGAAGAGGTGGAGCCGGATGCCACGAGGGGCTGTGAAGCAGAAGA
CCGGGGACTTCTCCGCCTCGCTGAAGCTGCAGGAAACCCTTCGAGGCATCCAAGTGTGGGG
CCCACCCTAATTACAGACCTTCCAAAAGATGACCGTGACCTTGAACCTTCCCTGGGGAGCCCTCC
TCTGACTGTGTGCTGGCGTCTCAAGCCTGAGTGCCTCCCGCTGGAGGAAGGGGAGTGCCACC
CTGTGTCCGTGGCCAGCACAGCGTACAACCTGACCCACACCTTCAGGGACCTTGGGGACTAC
TGCTTCAGCATCCGGGGCCGAGAATATCATCAGCAAGACACATCAGTACCACAAGATCCAGGT
GTGGCCCTCCAGAATCCAGCCGGCTGTCTTTGCTTTCCCATGTGCTACACTTATCACTGTGA
TGTTGGCCTTCATCATGTACATGACCCTGCGGAATGCCACTCAGCAAAAGGACATGGTGGAG
AACCCGGAGCCACCCTCTGGGGTCAAGGTGCTGCTGCCAGATGTGCTGTGGGCCTTTCTTGCT
GGAGACTCCATCTGAGTACCTGGAAATTGTTCTGTGAGAACACGGGGCTGCTCCCGCCCCCTCT
ATAAGTCTGTCAAACTTACACCGTG**TGA**GCACTCCCCCTCCCCACCCCATCTCAGTGTTAA
CTGACTGCTGACTTGGAGTTTCCAGCAGGGTGGTGTGCACCACTGACCAGGAGGGGTTTATT
TGCGTGGGGCTGTTGGCCTGGATCATCCATCCATCTGTACAGTTCAGCCACTGCCACAAGCC
CCTCCCTCTCTGTACCCCTGACCCACAGCCATTACCCATCTGTACAGTCCAGCCACTGACA
TAAGCCCCACTCGGTTACCACCCCTTGACCCCTACCTTTGAAGAGGCTTCGTGCAGGACT
TTGATGCTTGGGGTGTTCCGTGTTGACTCCTAGGTGGGCCTGGCTGCCCACTGCCCATTCCT
CTCATATTGGCACATCTGCTGTCCATTGGGGGTTCTCAGTTTCCCTCCCCCAGACAGCCCTAC
CTGTGCCAGAGAGCTAGAAAGAAGGTCATAAAGGGTTAAAAATCCATAACTAAAGGTTGTAC
ACATAGATGGGCACACTCACAGAGAGAAGTGTGCATGTACACACACCACACACACACACACA
CACACACACACAGAAATATAAACACATGCGTCACATGGGCATTTTCAAGATGATCAGCTCTGTA
TCTGGTTAAGTCGGTTGCTGGGATGCACCCTGCACTAGAGCTGAAAGGAAATTTGACCTCCA
AGCAGCCCTGACAGGTTCTGGGGCCCGGGCCCTCCCTTTGTGCTTTGTCTCTGCAGTTCTTGC
GCCCTTTATAAGGCCATCCTAGTCCCTGCTGGCTGGCAGGGGCTGGATGGGGGGCAGGACT
AATACTGAGTGATTGCAGAGTGCTTTATAAATATCACCTTATTTTATCGAAACCCATCTGTG
AAACTTTCACTGAGGAAAAGGCCTTGACGCGGTAGAAGAGGTTGAGTCAAGGCCGGGCGCGG
TGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGGCCGAGGCGGGTGGATCACGAGATCAGGA
GATCGAGACCACCTGGCTAACACGGTGAAACCCCGTCTCTACTAAAAAATAACAAAAGTT
AGCCGGGCGTGGTGGTGGTGCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATG
GTGCGAACCCGGGAGGCGGAGCTTGACAGTGAGCCAGATGGCGCCACTGCACTCCAGCCTGA
GTGACAGAGCGAGACTCTGTCTCCA

163/330

FIGURE 163

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLA
LPADAHLYRFHWIHTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGF
VVLPITEFLVGDLVVTQNTSLPWPSYLTKTVLKVSFLLHDPSNFLKTALFLYSWDFGDGTQ
MVTEDSVVYYNYSIIGTFTVKLKVVAEWEEVEPDATRAVKQKTGDFSASLKLQETLRGIQVL
GPTLIQTFQKMTVTLNFLGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYNLTHTFRDPGD
YCFSIRAENIISKTHQYHKIQVWPSRIQPAVFAFPCATLITVMLAFIMYMTLRNATQQKDMV
ENPEPPSGVRCCCQMCCGPFLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

Important features of the protein:**Signal peptide:**

amino acids 1-24

Transmembrane domain:

amino acids 339-362

N-glycosylation sites.

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367

165/330

FIGURE 165

MALSSQIWAACLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTH
FPICIFCCGCCHRSKCGMCCKT

166/330

FIGURE 166

CTGTCAGGAAGGACCATCTGAAGGCTGCAATTTGTTCTTAGGGAGGCAGGTGCTGGCCTGGC
CTGGATCTTCCACC**ATG**TTCCCTGTTGCTGCCTTTTGATAGCCTGATTGTCAACCTTCTGGGC
ATCTCCCTGACTGTCCTCTTCACCCTCCTTCTCGTTTTTCATCATAGTGCCAGCCATTTTTTG
AGTCTCCTTTGGTATCCGCAAACCTCTACATGAAAAGTCTGTTAAAAATCTTTGCGTGGGCTA
CCTTGAGAATGGAGCGAGGAGCCAAGGAGAAGAACCACCAGCTTTACAAGCCCTACACCAAC
GGAATCATTGCAAAGGATCCCACCTTCACTAGAAGAAGAGATCAAAGAGATTTCGTCTGAAGTGG
TAGTAGTAAGGCTCTGGACAACACTCCAGAGTTTCGAGCTCTCTGACATTTTCTACTTTTGCC
GGAAAGGAATGGAGACCATTATGGATGATGAGGTGACAAAGAGATTCTCAGCAGAAGAACTG
GAGTCCTGGAACCTGCTGAGCAGAACCAATTATAACTTCCAGTACATCAGCCTTCGGCTCAC
GGTCCTGTGGGGGTTAGGAGTGCTGATTTCGGTACTGCTTTCTGCTGCCGCTCAGGATAGCAC
TGGCTTTCACAGGGATTAGCCTTCTGGTGGTGGGCACAACTGTGGTGGGATACTTGCCAAAT
GGGAGGTTTAAGGAATTCATGAGTAAACATGTTCACTTAATGTGTTACCGGATCTGCGTGCG
AGCGCTGACAGCCATCATCACCTACCATGACAGGGAAAACAGACCAAGAAATGGTGGCATCT
GTGTGGCCAATCATACCTCACCGATCGATGTGATCATCTTGGCCAGCGATGGCTATTATGCC
ATGGTGGGTCAAGTGCACGGGGGACTCATGGGTGTGATTTCAGAGAGCCATGGTGAAGGCCTG
CCCACACGTCTGGTTTGAGCGCTCGGAAGTGAAGGATCGCCACCTGGTGGCTAAGAGACTGA
CTGAACATGTGCAAGATAAAAGCAAGCTGCCTATCCTCATCTTCCCAGAAGGAACCTGCATC
AATAATACATCGGTGATGATGTTCAAAAAGGGAAGTTTTTGAAATTGGAGCCACAGTTTACCC
TGTTGCTATCAAGTATGACCCTCAATTTGGCGATGCCTTCTGGAACAGCAGCAAATACGGGA
TGGTGACGTACCTGCTGCGAATGATGACCAGCTGGGCCATTGTCTGCAGCGTGTGGTACCTG
CCTCCCATGACTAGAGAGGCAGATGAAGATGCTGTCCAGTTTGCGAATAGGGTGAAATCTGC
CATTGCCAGGCAGGGAGGACTTGTGGACCTGCTGTGGGATGGGGGCTGAAGAGGGAGAAGG
TGAAGGACACGTTCAAGGAGGAGCAGCAGAAGCTGTACAGCAAGATGATCGTGGGGAACCAC
AAGGACAGGAGCCGCTCC**TGA**GCCTGCCTCCAGCTGGCTGGGGCCACCGTGCGGGGTGCCAA
CGGGCTCAGAGCTGGAGTTGCCGCCGCCGCCCCCACTGCTGTGTCCTTTCCAGACTCCAGGG
CTCCCCGGGCTGCTCTGGATCCCAGGACTCCGGCTTTCGCCGAGCCGCAGCGGGATCCCTGT
GCACCCGGCGCAGCCTACCCTTGGTGGTCTAAACGGATGCTGCTGGGTGTTGCGACCCAGGA
CGAGATGCCTTGTTTCTTTTACAATAAGTCGTTGGAGGAATGCCATTAAAGTGAACCTCCCA
CCTTTGCACGCTGTGCGGGCTGAGTGGTTGGGGAGATGTGGCCATGGTCTTGTGCTAGAGAT
GGCGGTACAAGAGTCTGTTATGCAAGCCCGTGTGCCAGGGATGTGCTGGGGGCGGCCACCCG
CTCTCCAGGAAAGGCACAGCTGAGGCACTGTGGCTGGCTTCGGCCTCAACATCGCCCCCAGC
CTTGAGCTCTGCAGACATGATAGGAAGGAAACTGTCATCTGCAGGGGCTTTCAGCAAAATG
AAGGGTTAGATTTTTTATGCTGCTGCTGATGGGGTTACTAAAGGGAGGGGAAGAGGCCAGGTG
GGCCGCTGACTGGGCCATGGGGAGAACGTGTGTTCTGACTCCAGGCTAACCCTGAACTCCCC
ATGTGATGCGCGCTTTGTTGAATGTGTGTCTCGGTTTCCCCATCTGTAATATGAGTCGGGGG
GAATGGTGGTGATTCCCTACCTCACAGGGCTGTTGTGGGGATTAAAGTGCTGCGGGTGAGTGA
AGGACACATCACGTTCAAGTACAGGCCACAAAACGGGGCACGGCAGGCCTGAG
CTCAGAGCTGCTGCACTGGGCTTTGGATTTGTTCTTGTGAGTAAATAAACTGGCTGGTGAA
TGA

167/330

FIGURE 167

MFLLLPFDSLIVNLLGISLTVLFTLLLVFIIIVPAIFGVVSFGIRKLYMKSLLKIFAWATLRME
RGAKEKNHQLYKPYTNGIIAKDPTSLEEEIKEIRRS GSSKALDNTPEFELSDIFYFCRKGME
TIMDDEVTKRFSAAEELESWNLLSRTNYNFQYISLRRLTVLWGLGVLIRYCFLLPLRIALFTG
ISLLVVGTTVVGYPNGRFKEFMSKHVHLMCYRICVRALTAIITYHDRENRP RNGGICVANH
TSPIDVILASDGYAMVGQVHGGLMGVIQ RAMVKACPHVWFERSEVKDRHLVAKRLTEHVQ
DKSKLPILIFPEGTCINNTSVMMFKKGSFEIGATVYPVAIKYDPQFGDAFWNSSKYGMV TYL
LRMMTSWAIVCSVWYLP PMTREADEDAVQFANRVKSAIARQGGLVDLLWDGGLKREKVKDTF
KEEQQKLYSKMIVGNHKDRSRS

168/330

FIGURE 168

GCCCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGCCCTCA
CGTCTCCTCCAGGG**ATG**GCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCAC
ACCTGGCAGGCCCAGGCTGTTCCCACCATCCTGCCCCCTGGGCCTGGCTCCAGACACCTTTGA
CGATACCTATGTGGGTTGTGCAGAGGAGATGGAGGAGAAGGCAGCCCCCTGCTAAAGGAGG
AAATGGCCCCACCATGCCCTGCTGCGGGAATCCTGGGAGGCAGCCCAGGAGACCTGGGAGGAC
AAGCGTCGAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGT
CTACACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAG
GCTCCCGGGAGCTCTACATGAGGCACTTTCCCTTCAAGGCCCTGCATTTCTACCTGATCCGG
GCCCTGCAGCTGCTGCGAGGCAGTGGGGGCTGCAGCAGGGGACCTGGGGAGGTGGTGTTCG
AGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGACTCTGTCCGCTTGGGCCAGT
TTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCACAGATTTGGGGAGAAGAGGCGGGGCTGT
GTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTTGCCCCCTG
GAAGACTCTGCTCTTGCCCCCTGGAGAGTTCCAGCTCTCAGGGGTTGGGCCC**TGA**AAGTCCA
ACATCTGCCACTTAGGAGCCCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAG
CAGCCTTGAGAAGCAAGAACATGGTTCCGGACCCAGCCCTAGCAGCCTTCTCCCCAACCAGG
ATGTTGGCCTGGGGAGGCCACAGCAGGGCTGAGGGAACTCTGCTATGTGATGGGGACTTCCT
GGGACAAGCAAGGAAAGTACTGAGGCAGCCACTTGATTGAACGGTGTTGCAATGTGGAGACA
TGGAGTTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGGA

169/330

FIGURE 169

MALAALMIALGSLGLHTWQAQAVPTILPLGLAPDTFDDTYVGCAEEMEEKAAPLLKEEMAHH
ALLRESWEAAQETWEDKRRGLTLPPGFKAQNGIAIMVYTNSNTLYWELNQAVRTGGGSREL
YMRHFPPFKALHFYLIRALQLLRGSGGCSRGPGEVVFRGVGSLRFEPKRLGDSVRLGQFASS
LDKAVAHRFGEKRRGCVSAPGVQLGSQSEGASSLPPWKTLLLAPGEFQLSGVGP

170/330

FIGURE 170

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAAT**ATG**GCTGGTTCCCCAACATGCCTCA
CCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTG
GTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTC
TATTGTCTGGACCTTCAACACAACCCCTCTTGTACCATAACAGCCAGAAGGGGGCACTATCA
TAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAG
CTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACT
CCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAG
TCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATG
GAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTC
CCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAGAAAGTGATATGACCTTCATCT
GCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCCATCCTTGCCAGGAAGCTCTGT
GAAGGTGCTGCTGATGACCCAGATTCCCTCCATGGTCCTCCTGTGTCTCCTGTTGGTGCCCCCT
CCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTCTGAAGAGAGAGAGACAAGAAG
AGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCTGGGAAACTCCTAACATATGCCCCCAT
TCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAATAGAACAATCCTAAAGGAAGA
TCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGATGGAAAATCCCCACTCAC
TGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTATC**TAG**ACAGCAGTG
CACTCCCCTAAGTCTCTGCTCA

171/330

FIGURE 171

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVT
IQPEGGTIIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHV
YEHLSPKPKVTMGLQSNKNGTCVTNLTCCMEHGEEVDVIYTWKALGQAANESHNGSILPISWRW
GESDMTFICVARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLW
FLKRERQEEYIEEKKRVDICRETPNICPHSGENTHEYDTIPHTNRTILKEDPANTVYSTVEIP
KKMENPHSLLTMPDTPRLFAYENVI

172/330

FIGURE 172

CTGGTTCCCCAACATGCCTCACCCCTCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCC
TCTGGACCCCGTGAAAGAGCTGGTTCGGTTCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTC
CAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTCACCATAC
AGCCAGAAGGGGGCACTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCA
GATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGT
GGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACG
AGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCAATAAGAATGGCACCTGTGTG
ACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATACCTGGAAGGCCCT
GGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAG
AAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCC
ATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCCCTCCATGGTCCTCCT
GTGTCTCCTGTTGGTGCCCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTC
TGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCTGGGAA
ACTCCTAACATATGCCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAA
TAGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAA
AGATGGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAG
AATGTTATCTAGACAGCAGTGCCTCCCCTAAGTCTCTGCTCAAAAAAAAAAAAAAAAAAAAAA

173/330

FIGURE 173

GAAAGACGTGGTCCTGACAGACAGACAATCCTATTCCCTACCAAA**ATGA**AGATGCTGCTGCT
GCTGTGTTTGGGACTGACCCTAGTCTGTGTCCATGCAGAAGAAGCTAGTTCTACGGGAAGGA
ACTTTAATGTAGAAAAGATTAATGGGGAATGGCATACTATTATCCTGGCCTCTGACAAAAGA
GAAAAGATAGAAGAACATGGCAACTTTAGACTTTTTCTGGAGCAAATCCATGTCTTGGAGAA
TTCCTTAGTTCTTAAAGTCCATACTGTAAGAGATGAAGAGTGCTCCGAATTATCTATGGTTG
CTGACAAAACAGAAAAGGCTGGTGAATATTCTGTGACGTATGATGGATTCAATACATTTACT
ATACCTAAGACAGACTATGATAACTTTCTTATGGCTCACCTCATTAACGAAAAGGATGGGGA
AACCTTCCAGCTGATGGGGCTCTATGGCCGAGAACCAGATTTGAGTTCAGACATCAAGGAAA
GGTTTGCACAACTATGTGAGGAGCATGGAATCCTTAGAGAAAATATCATTGACCTATCCAAT
GCCAATCGCTGCCTCCAGGCCCGAGAATGAAGAATGGCCTGAGCCTCCAGTGTTGAGTGGAC
ACTTCTCACCAGGACTCCACCATCATCCCTTCCTATCCATACAGCATCCCCAGTATAAATTC
TGTGATCTGCATTCCATCCTGTCTCACTGAGAAGTCCAATTCAGTCTATCAACATGTTACC
TAGGATACCTCATCAAGAATCAAAGACTTCTTTAAATTTCTCTTTGATACACCCTTGACAAT
TTTTCATGAAATTATTCCTCTTCCTGTTCAATAAATGATTACCCTTGCACTTAA

174/330

FIGURE 174

MKMLLLLCGLTLVCVHAEAEASSTGRNFNVEKINGEWHTIILASDKREKIEEHG NFR LFLEQ
IHVLENSLV LKVHTVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLI
NEKDGETFQLMGLYGREPDLSSDIKERFAQLCEEHGILRENIIDLSNANRCLQARE

175/330

FIGURE 175

GGCTCGAGCGTTTCTGAGCCAGGGGTGACC**ATG**ACCTGCTGCGAAGGATGGACATCCTGCAA
TGGATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTG
TCAGCTTAGTTGAGGAAGACCAATTTTCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTTC
CCAGGAATTATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAG
AAAAAGAGCGTGCTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCA
CAGTCATTGGTGCTCTGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTC
ATGTGTAATTCTCCAAGCAACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGA
CATTCATCCAGAATCCTTCAACTTGCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTG
GTTTCAATAAACCCACCAGTAACGACACCATGGCGAGTGGCTGGAGAGC**AT**CTAGTTTCCAC
TTCGATTCTGAAGAAAACAAACATAGGCTTATCCACTTCTCAGTATTTTTAGGTCTATTGCT
TGTTGGAATTCTGGAGGTCCTGTTTGGGCTCAGTCAGATAGTCATCGGTTTCCTTGGCTGTC
TGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTGTG**TAG**TTTAATGGGAATAAAATGTAAGTA
TCAGTAGTTTGAAAAAAAAAAAA

176/330

FIGURE 176

MTCCEGWTSCNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMA
IPATTMSLTARKKRACCNNRTGMFLSSFFSVITVIGALYCMLISIQALLKGPLMCNSPSNSNA
NCEFSLKNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRL
IHFSVFLGLLLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

177/330

FIGURE 177

GTCGAATCCAAATCACTCATTGTGAAAGCTGAGCTCACAGCCGAATAAGCCACC**ATG**AGGCT
GTCAGTGTGTCTCCTGATGGTCTCGCTGGCCCTTTGCTGCTACCAGGCCCATGCTCTTGTCT
GCCCAGCTGTTGCTTCTGAGATCACAGTCTTCTTATTCTTAAGTGACGCTGCGGTAAACCTC
CAAGTTGCCAAACTTAATCCACCTCCAGAAGCTCTTGCAGCCAAGTTGGAAGTGAAGCACTG
CACCGATCAGATATCTTTTAAGAAACGACTCTCATTGAAAAAGTCCTGGTGGAAA**TAG**TGAA
AAAATGTGGTGTGTGACATGTAAAAATGCTCAACCTGGTTTCCAAAGTCTTTCAACGACACC
CTGATCTTCACTAAAAATTGTAAAGGTTTCAACACGTTGCTTTAATAAATCACTTGCCCTGC

178/330

FIGURE 178

MRLSVCLLMVSLALCCYQAHALVCPAVASEITVFLFLSDAAVNLQVAKLNPPPEALAAKLEV
KHCTDQISFKKRLSLKKSWWK

179/330

FIGURE 179

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGG
TGAAGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTG
GAGCAGATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTC
GAACTGTGACATGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGG
AAGCCAATGACCCATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAG
CTGAGCGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGG
CAAATGCAAATACAAGAGCAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCAC
TCATCACTCCAGGCTCTGCCACTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGA
AGCCTAACACTGGCCCCCAGCACCTCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTC
TCTCCAAGGGCAGGCTGTTAGGCCCTTTCTGATCAGGAGGCTTCTTTATGAATTAACTCG
CCCCACCACCCCCTCA

180/330

FIGURE 180

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCK
YKSSQKQHSPVPEKAIPLITPGSATTC

181/330

FIGURE 181

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGG**ATG**TCGCTGCTGAGCCTGCCC
TGGCTGGGCCTCAGACCGGTGGCAATGTCCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTC
CTGGCTACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCC
AGTGTTTTCCACAGCCCCCAAACGGAACTGGTTTTGGGGTCACCTGGGCCTGATCACTCCT
ACAGAGGAGGGCTTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGT
ATGGCTGGGTCCCATCATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCA
CCAATGCCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCCCTGAAGCCCTGG
CTGGGAGAAGGGATACTGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGAC
GCCCCGCTTCCATTTCAACATCCTGAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACA
TCATGCTTGACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAG
CACATCAGCCTCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACAGCCATTG
TCAGGAGAGGCCCAGTGAATATATTGCCACCATCTTGGAGCTCAGTGCCCTTGTAGAGAAAA
GAAGCCAGCATATCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGGCGGCGC
TTCCACAGGGCCTGCCGCCTGGTGCATGACTTCACAGACGCTGTCATCCGGGAGCGGCGTCG
CACCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAAAGCCAAGTCCAAGACTTTGG
ATTTCAATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGGCATTGTCAGATGAGGAT
ATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCCTCTC
CTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGC
AAGAGCTTCTGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCCC
TTCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCCAGCTCCCTTCATCTCCCG
ATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCAAAGGCATTACCTGCC
TCATCGATATTATAGGGGTCCATCACAACCCA~~ACT~~GTGTGGCCGGATCCTGAGGTCTACGAC
CCCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCCCTTTCTC
CGCAGGGCCCAGGAACTGCATCGGGCAGGCGTTGCCATGGCGGAGATGAAAGTGGTCCTGG
CGTTGATGCTGCTGCACTTCGGTTCCCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAA
TTGATCATGCGCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCA
GTGACTTTCTGACCCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAAA

182/330

FIGURE 182

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPQPPKRNWFWG
HLGLITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLF
IRFLKPWLGE GILLSGGDKWSRHRRLTPAFHFNILKSYITIFNKSANIMLDKWQH LASEGS
SRLDMFEHISLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYY
LSHDGRRFHRACRLVHDFTDAVIRERRRTLPTQGIDDFFKDKAKSKTLD FIDVLLLSKDEDG
KALSDEDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEW
DDLAQLPFLTMCVKESLR LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVW
PDPEVYDPFRFDPENSKGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHT
EPRRKLELIMRAEGGLWLRVEPLNVGLQ

183/330

FIGURE 183

CAACAGAAGCCAAGAAGGAAGCCGTCTATCTTGTGGCGATC**ATG**TATAAGCTGGCCTCCTGC
TGTTTGCTTTTCACAGGATTCTTAAATCCTCTCTTATCTCTTCCTCTCCTTGACTCCAGGGA
AATATCCTTTCAACTCTCAGCACCTCATGAAGACGCGCGCTTAACTCCGGAGGAGCTAGAAA
GAGCTTCCCTTCTACAGATATTGCCAGAGATGCTGGGTGCAGAAAGAGGGGATATTCTCAGG
AAAGCAGACTCAAGTACCAACATTTTTTAACCCAAGAGGAAATTTGAGAAAGTTTCAGGATTT
CTCTGGACAAGATCCTAACATTTTACTGAGTCATCTTTTGGCCAGAATCTGGAAACCATAACA
AGAAACGTGAGACTCCTGATTGCTTCTGGAAATACTGTGTC**TGA**AGTGAAATAAGCATCTGT
TAGTCAGCTCAGAAACACCCATCTTAGAATATGAAAAATAACACAATGCTTGATTTGAAAAC
AGTGTGGAGAAAACTAGGCAAACTACACCCTGTTTCATTGTTACCTGGAAAATAAATCCTCT
ATGTTTTGCACAAAAAAAAAAAAAAAAA

184/330

FIGURE 184

MYKLASCCLLFTGFLNPLLSLPLLDREISFQLSAPHEDARLTPEELERASLLQILPEMLGA
ERGDILRKADSSTNIFNPRGNLRKFQDFSGQDPNILLSHLLARIWKPYKKRETPDCFWKYCV

185/330

FIGURE 185

GAACATTTTTAGTTCCCAAGGAATGTACATCAGCCCCACGGAAGCTAGGCCACCTCTGGGAT
GGGGTTGCTGGTTTAAAACAAACGCCAGTCATCCTATATAAGGACCTGACAGCCACCAGGCA
CCACCTCCGCCAGGAAGTGCAGGCCCCACCTGTCTGCAACCCAGCTGAGGCC**ATG**CCCTCCCC
AGGGACCGTCTGCAGCCTCCTGCTCCTCGGCATGCTCTGGCTGGACTTGGCCATGGCAGGCT
CCAGCTTCCTGAGCCCTGAACACCAGAGAGTCCAGCAGAGAAAGGAGTCGAAGAAGCCACCA
GCCAAGCTGCAGCCCCGAGCTCTAGCAGGCTGGCTCCGCCCCGGAAGATGGAGGTCAAGCAGA
AGGGGCAGAGGATGAACTGGAAGTCCGGTTCAACGCCCCCTTTGATGTTGGAATCAAGCTGT
CAGGGGTTTCAGTACCAGCAGCACAGCCAGGCCCTGGGGAAGTTTCTTCAGGACATCCTCTGG
GAAGAGGCCAAAGAGGCCCCAGCCGACAAG**TGA**TCGCCACAAAGCCTTACTCACCTCTCTCT
AAGTTTAGAAGCGCTCATCTGGCTTTTCGCTTGCTTCTGCAGCAACTCCCACGACTGTTGTA
CAAGCTCAGGAGGCGAATAAATGTTCAAACCTGTA

186/330

FIGURE 186

MPSPGTVCSLLLLGMLWLDLAMAGSSFLSPEHQRVQQRKESKKPPAKLQPRALAGWLRPEDG
GQAEGAEDELEVRFNAPFDVGIKLSGVQYQQHSQALGKFLQDILWEEAKEAPADKO

187/330

FIGURE 187

CGGCCACAGCTGGCATGCTCTGCCTGATCGCCATCCTGCTGTATGTCCTCGTCCAGTACCTC
GTGAACCCCGGGGTGCTCCGCACGGACCCCAGATGTCAAGAAT**ATGA**ACACGTGGCTGCTGT
TCCTCCCCCTGTTCCCGGTGCAGGTGCAGACCCTGATAGTCGTGATCATCGGGATGCTCGTG
CTCCTGCTGGACTTTCTTGGCTTGGTGCACCTGGGCCAGCTGCTCATCTTCCACATCTACCT
GAGTATGTCCCCCACCCTAAGCCCCCGATCCCCCAAGGCTGGGTGGTCAGAGCTGCTCATC
TTACACCTCTACTTGAGTATGTCCCTAACCCTGAGCCCCCAGCCTGGGGCCAGAGTCTTT
GTCCCCCGTGTGCGCATGTGTT**CAGGGT**CAGCCTCTCCCAGAAGTGAGATCATGGACAAAAA
GGGCAAATCACAGGAAGAAATTAAATCCATGAGGACCCAGCAGGCCCCAGCAAGAAGCTGAAC
TCACGCCGAGACCTGCAGGAGTGGTGCCAGGTGCT**TGA**AGTAACAAGTTTAAATGTT**CAGA**
GACAATGGAATGGAATCTATTAGGCAAGAACAGGACATTATGAAATAAGGACAGGTGGACTT
CCAAAAACACAAGTAGAAATTCTAACAATGAAATATATTACAGGCAGGT**CACCCACT**AACCA
AACAACTGAAGCGAGAGCTGTGGTCTTGCTTGGTCTCACAGTGGGCACAGCGGTAGGCGGTC
AGTCATGTTGCTGAACGACGGAGGGTAAACTCCCCAGCCCCAAGAAAACCTGTGTTGGAAGT
AACAAACAACCTCCCTGCTCCTGGCACCAGCCGTTTTGGTCATGGTGGGCCAGCTGCAAAGCG
TCTTCCATTCTCTGGGCAGTGGTGGCCCCGAGGCTGTGGCCTCTCAGGGGGTTTCTGTGGAC
ACGGGCAGCAGAGTGTGTCCAGGCCAGCCCCAAGAATGCCCTGCTCCTGACAGCTTGGCCA
ACCCCTGGTCAGGGCAGAGGGAGTTGGGTGGGT**CAGGCTCT**GGGCTCACCTCCATCTCCAGA
GCATCCCCTGCCTGCAGTTGTGGCAAGAACGCCCAGCTCAGAATGAACACACCCCAACCAAGA
GCCTCCTTGTT**CATA**ACCACAGGTTACCCTACAAACCACTGTCCCCACACAACCCTGGGGAT
GTTTTTAAACACACACCTCTAACGCATATCTTACAGTCACTGTTGTCTTGCCTGAGGGTTGA
ATTTTTTTTTAATGAAAGTGCAATGAAAATCACTGGATTAAATCCTACGGACACAGAGCTGAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

188/330

FIGURE 188

MNTWLLFLPLFPVQVQTLIVVIIGMLVLLLDLGLVHLGQLLIFHIYLSMSPTLSPRSPQGW
VVRAAHLTPLLEYVPNPEPPTPGARVFVPRVRMCSGSASPRSEIMDKKGKSQEEIKSMRTQQ
AQQEAELTPRPAGVPGA

189/330

FIGURE 189

GGAGTGCAGATGGCATCCTTCGGTTCTTCCAGACAAGCTGCAAGACGCTGACC**ATG**GCCAAG
ATGGAGCTCTCGAAGGCCTTCTCTGGCCAGCGGACACTCCTATCTGCCATCCTCAGCATGCT
ATCACTCAGCTTCTCCACAACATCCCTGCTCAGCAACTACTGGTTTGTGGGCACACAGAAGG
TGCCCAAGCCCCTGTGCGAGAAAGGTCTGGCAGCCAAGTGCTTTGACATGCCAGTGTCCCTG
GATGGAGATACCAACACATCCACCCAGGAGGTGGTACAATACAACCTGGGAGACTGGGGATGA
CCGGTTCTCCTTCCGGAGCTTCCGGAGTGGCATGTGGCTATCCTGTGAGGAAACTGTGGAAG
AACCAGGGGAGAGGTGCCGAAGTTTCATTGAACTTACACCACCAGCCAAGAGAGGTGAGAAA
GGACTACTGGAATTTGCCACGTTGCAAGGCCCATGTCACCCCACTCTCCGATTTGGAGGGAA
GCGGTTGATGGAGAAGGCTTCCCTCCCCTCCCCTCCCCTTGGGGCTTTGTGGCAAAAATCCTA
TGGTTATCCCTGGGAACGCAGATCACCTACATCGGACTTCAATTCATCAGCTTCCTCCTGCT
ACTAACAGACTTGCTACTCACTGGGAACCCTGCCTGTGGGCTCAAACCTGAGCGCCTTTGCTG
CTGTTTCCTCTGTCCTGTCAGGTCTCCTGGGGATGGTGGCCACATGATGTATTCAACAAGTC
TTCCAAGCGACTGTCAACTTGGGTCCAGAAGACTGGAGACCACATGTTTGGAATTATGGCTG
GGCCTTCTACATGGCCTGGCTCTCCTTCACCTGCTGCATGGCGTCGGCTGTCACCACCTTCA
ACACGTACACCAGGATGGTGCTGGAGTTCAAGTGCAAGCA**TAG**TAAGAGCTTCAAGGAAAAC
CCGAAGTGCCTACCACATCACCATCAGTGTTTCCCTCGGCGGCTGTCAAGTGCAGCCCCCAC
CGTGGGTCCTTTGACCAGCTACCACCAGTATCATAATCAGCCCATCCACTCTGTCTCTGAGG
GAGTCGACTTCTACTCCGAGCTGCGGAACAAGGGATTTCAAAGAGGGGCCAGCCAGGAGCTG
AAAGAAGCAGTTAGGTCATCTGTAGAGGAAGAGCAGTGTTAGGAGTTAAGCGGGTTTGGGGA
GTAGGCTTGAGCCCTACCTTACACGTCTGCTGATTATCAACATGTGCTTAAGCCAACATCCG
TCTCTTGAGCATGGTTTTTTAGAGGCTACGAATAAGGCTATGAATAAGGGTTATCTTTAAGTC
CTAAGGGATTCTTGGGTGCCACTGCTCTCTTTTCTCTACAGCTCCATCTTGTTTCACCCAC
CCCACATCTCACACATCCAGAATTCCTTCTTTACTGATAGTTTCTGTGCCAGGTTCTGGGC
TAAACCATGGAGATAAAAAGAAGAGTAAAATACACTTCCCGACCTTAAGGATCTGAAA

190/330

FIGURE 190

MAKMELSKAFSGQRTLLSAILSMLSLSFSTTSLLSNYWFVGTQKVPKPLCEKGLAAKCFDMP
VSLDGD TNTSTQEVVQYNWETGDDRFSFRSFRSGMWLSCEETVEEPGERCRSFIELTPPAKR
GEKGLLEFATLQGPCHPTLRFGGKRLMEKASLPSPPLGLCGKNPMVIPGNADHLHRTSIHQL
PPATNRLATHWEPCLWAQTERLCCCFLCPVRSPGDGGPHDVFTSLPSPDCQLGSRRLTTCLE
LWLGLLHGLALLHLLHGVGCHHLQHVHQDGAGVQVQA

191/330

FIGURE 191

AACTGGAAGGAAAGAAAGAAAGGTCAGCTTTGGCCCAG**ATG**TGGTTACCCCTTGGTCTCCTG
TCTTTATGTCTTTCTCCTCTTCCTATTCTGTCATCTCCCTCACTTAAGTCTCAGGCCTGTCA
GCAGCTCCTGTGGACATTGCCATCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATG
TTATGGATGTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGT
GATCTCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTGAAATCTGGCATG
AGATGGCACAGGTGACCACGCAGAAGCCACCAGAATCTTGCCTGCCCTATTCTCCTCCCAA
GTCTGTTCTCTTATTGTCAACCTCAGCACAAACAGGCTGGCGCCAATGGCATTACAGAGAAAG
CAATCTGTGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGT
AGCCACCTCCCTGTGTCAGCCAGTATTAACATGTCCCCTTCCCCCTGCCCCGCCGTAGATTGAG
GACATTGCCCCCTGTGTGCCACCAAACCAGGACTTTCCCCTTGGCTTGGCATCCCTGGCTCT
CTCCTGGTACCCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTTACT
ATGGCGATGGCCATGATGTTACAATCCCCTTGCCTGAATAATCAAGTGGGAAGGGGAAGCA
GAGGGAAATGGGGCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCTGAGGAAAAACCAAA
GGGAAGCAACAGGAACTTCTGCAACTGGTTTTTATCGGAAAGATCATCCTGCCTGCAGATGC
TGTTGAAGGGGCACAAGAAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAA
ATAGAACAGTCTGCTGGGAGTCAGACCTGGAATTCTGATTCCAAACTCTTTATTACTTTGGG
AAGTCACTCAGCCTCCCCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGG
AACCAAGGAAACTAACAATGTAGGTTACTAGTGAATACCCCAATGGTTTCTCCAATTATGCC
CATGCCACCAAAACAATAAAACAAAATTCTCTAACACTGAAA

192/330

FIGURE 192

MWLPLGLLSLCLSPILSSPSLKSQACQQLLWTLPSPLVAFRANRTTYVMDVSTNQSGME
HRNHLCFCDLYDRATSPPLKCSLL

193/330

FIGURE 193

GTAGCGCGTCTTGGGTCTCCCGGCTGCCGCTGCTGCCGCCGCCGCTCGGGTTCGTGGAGCCAGGAGCGACGTCA
CCGCC**ATGGC**AGGCATCAAAGCTTTGATTAGTTTGTCTTTGGAGGAGCAATCGGACTGATGTTTTTGATGCTT
GGATGTGCCCTTCCAATATACAACAAATACTGGCCCCCTCTTTGTTCTATTTTTTTACATCCTTTCACCTATTCC
ATACTGCATAGCAAGAAGATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACCTGGCCATCTTTC
TTACAACGGGCATTGTCTGTCTCAGCTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGATTGAGTGGGGA
GCTTGTGCACTTGTCTCTCACAGGAAACACAGTCATCTTTGCAACTATACTAGGCTTTTTCTTGGTCTTTGGAAG
CAATGACGACTTCAGCTGGCAGCAGTGG**TGA**AAAGAAATTACTGAACTATTGTCAAATGGACTTCCTGTCAATTT
GTTGGCCATTACGCACACAGGAGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTA
GGTGCTCCCTTCTCACTTTTATTGTAAGCATACTATTTTCACAGAGACTTGCTGAAGGATTAAAAGGATTTTCT
CTTTTGAAAAGCTTGACTGATTTACACTTATCTATAGTATGCTTTTTGTGGTGTCTGCTGAATTTAAATAT
TTATGTGTTTTCTGTAGGTTGATTTTTTTGGGAATCAATATGCAATGTTAAACACTTTTTTAATGTAATCA
TTTGCATTGGTTAGGAATTCAGAATTCGCGCCGGCTCTATTACTGGTCAAGTACATCTTTTCTCTTAAATATT
TAGCCTCCATTATTACAAAAAATTATAAAAAATAAGTTTTCAGTCAGTCAGGATGACATCACTCCCAATGTTATG
CAGACATACAGACGGTTGGCATACTGTTATAGACTGTATACTCAGTGCAAATATAGCTGCATTTATACCTCAGAG
GGGCCAAGTGTTAATGCCCATGCCCTCCGTTAAGGGTTGTTGGTTTTACTGGTAGACAGATGTTTTGTGGATTG
AAAATTATTTTATGGAATTGCTACAGAGGAGTGCTTTTCTCTCAATTGTTAGAAGAATTTATGTTAACTTTA
AGGTAAGGGTGTAACAAACATTTTTGAGATAAGGTTTTTATTTATGTTTATTATTGTTAGAGTGAGTTGCAATGT
GGGAAGAAATGACATTGAAATTCAGTTTTTGAATCCTGTTTCTATTTATAAGTGAAATTTGTGATCTCCTATC
AACCTTTTCATGTTTTACCCTGTTAAATGGACATACATGGAACCACTACTGATGAGGGACAGTTGTATGTTTGC
ATCATATATGCCAGAAAACCTTCCTCTGCTTCCTCCTTTTGACTTATTTGGTATGTTGTATATATTACATAAAA
TAACTTTTCAAATATAGTTTAATAACACTTAGAAGTGTTTACTTACCTGGAAAATAATTGCTATGCCGTACATT
CAGAGTGCCCCCTCCCCTGCAAGGCCTTGCCATGATTAACAAGTAACCTGTTAGTCTTACAGATAATTCATGCA
TTAACAGTTTAAGATTTAGACCATGGTAATAGTAGTTCTTATTCTCTAAGGTTATATCATATGTAATTTAAAG
TATTTTAAAGACAAGTTTCTGTATACCTCTGAACGTGTTTTGATTTTGAGTTCATCATGATAGATCTGCTGTTT
CCTTATAAAAGGCATTTGTTGTGTGAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACAT
ACCTGACCAAAAAATTTCCAGTAACCAGGCATGATCAATTTATAGTGGTTCGTTTACATCTAATAATTATCAGGA
CTTTTTTCAGGAGTGGGTTATAAAAAACATTCAAGTTGGTCTGACAGTATTTTGTTAAGGATATTTGTTTGTATG
TTTATTCAGTATACTTACATAAAAAATTATTTGCCATCAGCCAAACTCAGTAATCATGACAGCTGTCTGTTGT
TTTATGAAGTTTATTTCTCAAGAAAAATGGGAATAAATTTGGGATTTGTTTCAGCTTTTTTACTAAAGATGCCTAA
AGCCACAGGTTTTATTGCCTAACTTAAGCCATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCG
GCGTGTGGCTGGAGCCTTCCCCTGGAGGCTGAAAGTGGCTTGTGGTATTATAATGTTTCAGATTTCAAGAGGAA
GGTGCAAGGTACACATGAGTTAGAGAGCTGGTGAGACAGTTGGGAACCTTTGTGCTTGTGATCTACTGGACTTT
TTTTTTCAGGAAGTGCATTCTCTGGTCCCTCCCTATTTTCTGTTCTGGATGTCAGTGCAGTGCAGTGCCTACTG
TTTTATCCACTTGGCCACAGACTTTTTCTAACAGCTGCGTATTATTTCTATATACTAATTGCATTGGCAGCATT
GTGTCTTTGACCTTGATACTAGCTTGACATAGTGCTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAAT
TTTCCATAGAATATGCACTGATACAACATTACCATTCTTCTATGGAAAGAAAACCTTTTGATGATGAAACAATAA
AGATTTTAAATATCTATTTTAAAAA

194/330

FIGURE 194

MAGIKALISLSFGGAIGLMFLMLGCALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTDAM
SNACKELAIFLTGTGIVVSAFGLPIVFARAHLEWGACALVLTGNTVIFATILGFFLVFGSND
DFSWQQW

FIGURE 195

[illegible]

196/330

FIGURE 196

MDFLLLGLCLYWLLRRPSGVVLCLLGACFQMLPAAPSGCPQLCRCEGRLLYCEALNLTEAPH
NLSGLLGLSLRYNSLSELRAGQFTGLMQLTWLYLDHNNHICSVQGDAFQKLRRVKELTLSSNQ
ITQLPNTTFRPMPNLRSDLSYNKLQALAPDLFHGLRKLTLHMRANAIQFVPVRIFQDCRS
LKFLDIGYNQLKSLARNSFAGLFKLTELHLEHNDLVKVNFAHFPRILSLHSLCLRRNKVAIV
VSSLDWVWNLEKMDLSGNEIEYMEPHVFETVPHLQSLQLDSNRLTYIEPRILNSWKSLSIT
LAGNLWDCGRNVCALASWLSNFQGRYDGNLQCASPEYAQGEDVLDAVYAFHLCEDGAEPTSG
HLLSAVTNRSDLGPPASSATTADGGEGQHDGTFEPATVALPGGEHAENAVQIHKVVTGTMA
LIFSFLIVVLVLYVSWKCFPASLRQLRQCFVTQRRKQKQKQTMHOMAAMSAQEYYVDYKPNH
IEGALVIINEYGSCTCHQQPARECEV

197/330

FIGURE 197

GTGCAAGGAGCCGAGGCGAG**ATG**GGCGTCCTGGGCCGGGTCCCTGCTGTGGCTGCAGCTCTGC
GCACTGACCCAGGCGGTCTCCAAACTCTGGGTCCCCAACACGGACTTCGACGTCGCAGCCAA
CTGGAGCCAGAACCGGACCCCGTGCGCCGGCGGCGCCGTTGAGTTCCCGGCGGACAAGATGG
TGTCAGTCCTGGTGCAAGAAGGTCACGCCGTCTCAGACATGCTCCTGCCGCTGGATGGGGAA
CTCGTCCTGGCTTCAGGAGCCGGATTTCGGCGTCTCAGACGTGGGCTCGCACCTGGACTGTGG
CGCGGGCGAACCTGCCGTCTTCCGCGACTCTGACCGCTTCTCCTGGCATGACCCGCACCTGT
GGCGCTCTGGGGACGAGGCACCTGGCCTCTTCTTCGTGGACGCCGAGCGCGTGCCCTGCCGC
CACGACGACGTCTTCTTTCCGCCTAGTGCCTCCTTCCGCGTGGGGCTCGGCCCTGGCGCTAG
CCCCGTGCGTGTCGCGAGCATCTCGGCTCTGGGCCGGACGTTACGCGCGACGAGGACCTGG
CTGTTTTCTGCGTCCCGCGCGGGCCGCCTACGCTTCCACGGGCCGGGCGCGC**TGA**GCGTG
GGCCCCGAGGACTGCGCGGACCCGTCGGGCTGCGTCTGCGGCAACGCGGAGGCGCAGCCGTG
GATCTGCGCGGCCCTGCTCCAGCCCCCT

198/330

FIGURE 198

MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLVQE
GHAVSDMLLPLDGELVLASGAGFGVSDVGSHLDCGAGEPAVFRDSDRFSWHDPHLWRSGDEA
PGLFFVDAERVPCRHDDVFFPPSASFRVGLGPGASPVRVRSISALGRTFTRDEDLAVFLASR
AGRLRFHGPGALSVGPEDCADPSGCVCGNAEAQPWICAALLQP

FIGURE 199

[illegible]

200/330

FIGURE 200

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIP
FARDAVKKCFVCLA

201/330

FIGURE 201

TTGAGCGCAGGTGAGCTCCTGCGCGTTCCGGGGGCGTTCCCTCCAGTCACCCCTCCCGCCGTTACCCGCGGCGCGC
CCGAGGGAGTCTCCTCCAGACCCCTCCCTCCCGTTGCTCCAACTAATACGGACTGAACGGATCGCTGCGAGGGT
GGGAGAGAAAATTAGGGGGAGAAAGGACAGAGAGAGCAACTACCATCCATAGCCAGATAGATTATCTTACACTG
AACTGATCAAGTACTTTGAAAATGACTTCGAAATTTATCTTGGTGTCTTCATACTTGCTGCACTGAGTCTTTC
AACCACCTTTTCTCTCCAAGTACAGCAGCAAAAGGTTCTACTAGTTTCTTTTGATGGATTCCGTTGGGATTACT
TATATAAAGTTCCAACGCCCCATTTTCATTATATTATGAAATATGGTGTTCACGTGAAGCAAGTTACTAATGTT
TTTATTACAAAACCTACCCTAACCATTTATCTTTGGTAACTGGCCTCTTTCAGAGAAATCATGGGATTGTTGC
AAATGATATGTTTGATCCTATTTCGGAACAAATCTTCTCCTTGGATCACATGAATATTTATGATTCCAAGTTTT
GGGAAGAAGCGACACCAATATGGATCACAAACCAGAGGGCAGGACATACTAGTGGTGCAGCCATGTGGCCCCGA
ACAGATGTAAAAATACATAAGCGCTTTCCTACTCATTACATGCCTTACAATGAGTCAGTTTCATTTGAAGATAG
AGTTGCCAAAATTGTTGAATGGTTTACGTCAAAAGAGCCCATAAATCTTGGTCTTCTCTATTGGGAAGACCCTG
ATGACATGGGCCACCATTTGGGACCTGACAGTCCGCTCATGGGGCCTGTCAATTCAGATATTGACAAGAAGTTA
GGATATCTCATACAAATGCTGAAAAAGGCAAGTTGTGGAACACTCTGAACCTAATCATCACAAGTGATCATGG
AATGACGCAGTGCTCTGAGGAAAGGTTAATAGAACTTGACCAGTACCTGGATAAAGACCCTATACCCTGATTG
ATCAATCTCCAGTAGCAGCCATCTTGCCAAAAGAAGGTAAATTTGATGAAGTCTATGAAGCACTAACTCACGCT
CATCCTAATCTTACTGTTTACAAAAAAGAAGACGTTCCAGAAAGGTGGCATTACAAATACAACAGTCGAATTCA
ACCAATCATAGCAGTGGCTGATGAAGGGTGGCACATTTTACAGAATAAGTCAGATGACTTTCTGTTAGGCAACC
ACGGTTACGATAATGCGTTAGCAGATATGCATCCAATATTTTTAGCCCATGGTCTTGCCTTCAGAAAGAATTTT
TCAAAAGAAGCCATGAACCTCCACAGATTTGTACCCACTACTATGCCACCTCCTCAATATCACTGCCATGCCACA
CAATGGATCATTTCTGGAATGTCCAGGATCTGCTCAATTCAGCAATGCCAAGGGTGGTCCCTTATACACAGAGTA
CTATACTCCTCCCTGGTAGTGTTAAACCAGCAGAATATGACCAAGAGGGGTCAATCCCTTATTTCATAGGGGT
TCTCTTGGCAGCATTATAGTGATTGTATTTTTTGTAAATTTTCATTAAGCATTTAATTCACAGTCAAATACCTGC
CTTACAAGATATGCATGCTGAAATAGCTCAACCATTATTACAAGCCTAATGTTACTTTGAAGTGGATTTGCATA
TTGAAGTGGAGATTCCATAATTATGTCAGTGTTTTAAAGGTTTTCAAATTTCTGGGAAACCAGTTCCAAACATCTGC
AGAAACCATTAAGCAGTTACATATTTAGGTATACACACACACACACACACATACACACACACGGACCAAA
ATACTTACACCTGCAAAGGAATAAAGATGTGAGAGTATGTCTCCATTGTTCACTGTAGCATAGGGATAGATAAG
ATCCTGCTTTATTTGGACTTGGCGCAGATAATGTATATATTTAGCAACTTTGCACTATGTAAAGTACCTTATAT
ATTGCACTTTTAAATTTCTCTCCTGATGGGTACTTTAATTTGAAATGCACCTTATGGACAGTTATGTCTTATAAC
TTGATTGAAAATGACAACCTTTTTGCACCCATGTCACAGAATACTTGTTACGCATTGTTCAAACCTGAAGGAAATT
TCTAATAATCCCGAATAATGAACATAGAAATCTATCTCCATAAATTGAGAGAAGAAGAAGGTGATAAGTGTTGA
AAATTAAATGTGATAACCTTTGAACCTTGAATTTTGGAGATGTATTCCCAACAGCAGAATGCAACTGTGGGCAT
TTCTTGTCTTATTTCTTTCCAGAGAACGTGGTTTTTCATTTATTTTTCCCTCAAAAGAGAGTCAAATACTGACAG
ATTCGTTCTAAATATATTGTTTCTGTCAAAAATTATTGTGATTTCTGATGAGTCATATTACTGTGATTTTCA
TAATAATGAAGACACCATGAATATACTTTTCTCTATATAGTTTCAGCAATGGCCTGAATAGAAGCAACCAGGCA
CCATCTCAGCAATGTTTTCTCTTGTGTTGTAATTATTTGCTCCTTTGAAAATTAAATCACTATTAATTACATTAA
AAATCAAATTGGATAAAAAAAAAAAAAAAAAAAAAA

202/330

FIGURE 202

MTSKFILVSFILAALSLSTTFSLQLDQQKVLLVSFDGFRWDYLYKVPTPHFHYIMKYGVHVK
QVTNVFITKTYPNHYTLVTGLFAENHGIVANDMFDPIRNKSFSLDHMNIYDSKFWEEATPIW
ITNQ RAGHTSGAAMWPGTDVKIHKRFPTHYMPYNESVSFEDRVAKIVEWFTSKEPINLGLLY
WEDPDDMGHHLGPDSPLMGPVISDIDKKLGYLIQMLKKAKLWNTLNLIITSDHGMTQCSEER
LIELDQYLDKDHYTELIDQSPVAAAILPKEGKFDEVYEALTHAHPNLTVYKKEDVPERWHYKYN
SRIQPIIAVADEGWHILQNKSDDFLLGNHGYDNALADMHPIFLAHGPAFRKNFSKEAMNSTD
LYPLLCHLLNITAMPHNGSFWNVQDLLNSAMPRVVPYTQSTILLPGSVKPAEYDQEGSYPYF
IGVSLGSIIVIVFFVIFIKHLIHSQIPALQDMHAEIAQPLLQA

Signal Peptide:

amino acids 1-22

Transmembrane Domain:

amino acids 429-452

N-glycosylation sites:

amino acids 101-104, 158-161, 292-295, 329-332, 362-365, 369-372, 382-385, 389-392

Somatomedin B Domain:

amino acids 69-85

Sulfatase protein Region:

amino acids 212-241

203/330

FIGURE 203

GGATTTTTGTGATCCGCGATTTCGCTCCACGGGCGGGACCTTTGTAACTGCGGGAGGCCCAG
GACAGGCCCACCCTGCGGGGCGGGAGGCAGCCGGGGTGAGGGAGGTGAAGAAACCAAGACGC
AGAGAGGCCAAGCCCCTTGCCCTGGGGTCACACAGCCAAAGGAGGCAGAGCCAGAACTCACAA
CCAGATCCAGAGGCAACAGGGAC**ATG**GCCACCTGGGACGAAAAGGCAGTCACCCGCAGGGCC
AAGGTGGCTCCCGCTGAGAGGATGAGCAAGTTCTTAAGGCACTTCACGGTCGTGGGAGACGA
CTACCATGCCTGGAACATCAACTACAAGAAATGGGAGAATGAAGAGGAGGAGGAGGAGGAGG
AGCAGCCACCACCCACACCAGTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCTGACGTTGCC
CCTGCCCCCTGGCCCCGCACCCAGGGCCCCCCTTGACTTCAGGGGCATGTTGAGGAACTGTT
CAGCTCCCACAGGTTTCAGGTCATCATCATCTGCTTGGTGGTTCTGGATGCCCTCCTGGTGC
TTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCGACAAGAATAACTATGCTGCCATG
GTATTCCACTACATGAGCATCACCATCTTGGTCTTTTTTATGATGGAGATCATCTTTAAATT
ATTTGTCTTCCGCCTGAGTTCTTTCACCACAAGTTTGAGATCCTGGATGCCCGTCGTGGTGG
TGGTCTCATTCATCCTGGACATTGTCCTCCTGTTCCAGGAGCACCAGTTTGAGGCTCTGGGC
CTGCTGATTCTGCTCCGGCTGTGGCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGT
TAAGACACGTTCAGAACGGCAACTCTTAAGGTTAAAACAGATGAATGTACAATTGGCCGCCA
AGATTCAACACCTTGAGTTCAGCTGCTCTGAGAAGCCCCCTGGACT**TGA**TGAGTTTGCTGTATC
AACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTACTCT
CACACAGCCACCGTGAAAGTCCTGGAGTAAAATGTGCTGTGTACAGAAGAGAGAGAAGGAAG
CAGGCTGGCATGTTCACTGGGCTGGTGTACGACAGAGAACCTGACAGTCACTGGCCAGTTA
TCACTTCAGATTACAAATCACACAGAGCATCTGCCTGTTTTCAATCACAAGAGAACAAAACC
AAAATCTATAAAGATATTCTGAAAATATGACAGAATTTGACAAATAAAAGCATAAACGTGTA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

204/330

FIGURE 204

MATWDEKAVTRRAKVAPAERMSKFLRHFTVVGDDYHAWNINYYKKWENEEEEEEEEQPPPTPV
SGEEGRAAAPDVAPAPGPAPRAPLDFRGMLRKLFSSSHRFQVIIICLVVLDALLVLAELIDL
KIIQPDKNNYAAMVFHYMSITILVFFMMEIIFKLFVFRLLSSFTTSLRSWMPVVVVVSFILD
VLLFQEHQFEALGLLILLRLWRVARIINGIIISVKTRSERQLRLKQMNVQLAAKIQHLEFS
CSEKPLD

FIGURE 205

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCGAGTGGAGCACCCAGCAGGCCGCCAAC**AT**
GCTCTGTCTGTGCCTGTACGTGCCGGTCAATCGGGGAAGCCCAGACCGAGTTCCAGTACTTTG
AGTCGAAGGGGCTCCCTGCCGAGCTGAAGTCCATTTTCAAGCTCAGTGTCTTCATCCCCCTCC
CAGGAATTCCTCCACCTACCCGCCAGTGGGAAGCAGAAATTTGTACAAGCTGGAGATAAGGACCT
TGATGGGCGAGCTAGACTTTGAAGAATTTGTCCATTATCTCCAAGATCATGAGAAGAAGCTGA
GGCTGGTGTTTAAGATTTTGGACAAAAAGAATGATGGACGCATTGACGCGCAGGAGATCATG
CAGTCCCTGCGGGACTTGGGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCTCAAGAG
CATGGATAAAAAACGGCACGATGACCATCGACTGGAACGAGTGGAGAGACTACCACCTCCTCC
ACCCCGTGGAAAAACATCCCCGAGATCATCTCTACTTGAAGCATTCACAGATCTTTGATGTG
GGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGAGGAGAGGCAGACGGGGATGTGGTG
GAGACACCTGGTGGCAGGAGGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCCCCCTGG
ACAGGCTCAAGGTGCTCATGCAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGT
GGCTTCACTCAGATGATTCGAGAAGGAGGGGCCAGGTCACTCTGGCGGGGCAATGGCATCAA
CGTCTCTAAAATTTGCCCCGAATCAGCCATCAAATTCATGGCCTATGAGCAGATCAAGCGCC
TTGTGGTAGTGACCCAGGAGACTCTGAGGATTCACAGAGAGGCTTGTGGCAGGGTCTTGGCA
GGGGCCATCGCCCAGAGCAGCATCTACCCAATGGAGGTCTGAAGACCCGGATGGCGCTGCG
GAAGACAGGCCAGTACTCAGGAATGCTGGACTGCGCCAGGAGGATCCTGGCCAGAGAGGGGG
TGGCCGCCTTCTACAAAGGCTATGTCCCCAACATGCTGGGCATCATCCCCATATGCCGGCATC
GACCTTGCAGTCTACGAGACGCTCAAGAATGCCTGGCTGCAGCATATGCAGTGAACAGCGC
GGACCCCGCGGTGTTTGTGCTCTGGCCTGTGGCACCATGTCCAGTACCTGTGGCCAGCTGG
CCAGCTACCCCTTGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTCTATTGAGGGCGCT
CCGGAGGTGACCATGAGCAGCCTCTTCAAACATATCCTGCGGACCGAGGGGGCCTTCGGGCT
GTACAGGGGGCTGGCCCCCACTTCATGAAGGTCACTCCAGCTGTGAGCATCAGCTACGTGG
TCTACGGAACACTGAAGATCACCTGGGCGTGCAGTCCGCG**TG**ACGGGGGAGGGCCGCCCG
CTAGTGGAATCGCTGATCCTGGGCGCAGCCTGGGGTGTGCAGCCATCTCATTCTGTGAATG
TGCCAACACTAAGCTGTCTCGAGCCAAGCTGTGAAAACCCTAGACGCACCCGCAGGGAGGGT
GGGGAGAGCTGGCAGGCCCAGGGCTTGTCTGCTGACCCAGCAGACCCTCCTGTTGGTTCC
AGCGAAGACCACAGGCATTCCTTAGGGTCCAGGGTCAAGCAGGCTCCGGGCTCATATGTGTAA
GGACAGGACATTTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCTTGGAGGCCGGCTTAGT
TCTTCCATTTACCCCTTGCAGCCAGCTGTGGCCACGGCCCTGCCCTCTGGTCTGCCCTGC
ATCTCCCTGTGCCCTCTTGCTGCCTGCCTGTCTGCTGAGGTAAGGTGGGAGGAGGGCTACAG
CCCACATCCCACCCCTCGTCCAATCCCATAATCCATGATGAAAGGTGAGGTACAGTGGCCT
CCCAGGCTGACTTCCCAACCTACAGCATGACGCCAATCTGGCTGTGAAGGAAGGAAAG
GATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGCTCTCGGGCATGCT
TGGGAGTGCAGGGGGCTCGGGCTGCCTGGCCTGGCTGCACAGAAGGCAAGTGTGGGGCTCA
TGGTGTCTCTGAGCTGGCCTGGACCCTGTCAGGATGGGCCCCACCTCAGAACCAAATCACTG
TCCCCACTGTGGGATGAGGGCAGTGGAGCACCATTGTTGAGGGCGAAGGGCAGAGCGTTTGT
GTGTTCTGGGGAGGGAAGGAAAAGGTGTTGAGGCTTAATTTATGGAATGTTGGGAAAAGGG
TTTTGTCCAGAAGGACAAGCCGGAACAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTTCCTGTCCAACC
CCAGCAGGGGGCGCAGCGGGACCAGCCCCACATTCACATTGTGTCACTGCTTGGAACCTATTT
ATTTTGTATTTATTTGAACAGAGTTATGTCTTAACATATTTTATAGATTTTGTTTAATTAATA
GCTTGTCACTTTTCAAGTTCACTTTTATTTATATTTATGTTCACTGTTGATTTGTACCTTCCC
AAGCCCGCCAGTGGGATGGGAGGAGGAGGAGAAGGGGGCCTTGGGGCGCTGCAGTCACAT
CTGTCCAGAGAAATTCCTTTTGGGACTGGAGGCAGAAAAGCGGCCAGAAGGCAGCAGCCCTG
GCTCCTTTTCTTTTGGCAGGTTGGGGAAGGGCTTGGCCCCAGCCTTAGGATTTACAGGGTTTGA
CTGGGGGGCGTGGAGAGGAGAGGGAGGAACCTCAATAACCTTGAAGGTGGAATCCAGTTATTTT
CTGCGCTGCGAGGGTTTCTTTATTTCACTCTTTTCTGAATGTCAAGGCAGTGGAGGTGCTCT
CACTGTGAATTTGTGGTGGGCGGGGCTGGAGGAGAGGGTGGGGGCTGCGCTCCGTCCCTCC
CAGCCTTCTGCTGCCCTTGCTTAACAATGCCGGGCCAACTGGCGACCTCACGGTTGCACTTCC
ATTCCACCAGAATGACCTGATGAGGAAATCTTCAATAGGATGCAAAGATCAATGCAAAAATT
GTTATATATGAACATATAACTGGAGTCGTCAAAAAGCAAATTAAGAAAGAATTGGACGTTAG
AAGTTGTCAATTTAAAGCAGCCTTCTAATAAAGTTGTTTCAAAGCTGAAAAAATAAAAAA
AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA

206/330

FIGURE 206

MLCLCLYVPVIGEAQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKD
LDGQLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQAEEKILK
SMDKNGTMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTPDEFTVEERQTGMW
WRHLVAGGGAGAVSRTCTAPLDRLKVLMOVHASRSNNMGIVGGFTQMIREGGARSLWRGNGI
NVLKIAPESAIFMAYEQIKRLVGSDQETLRIHERLVAGSLAGAIQAQSSIYPMEVLKTRMAL
RKTGQYSGMLDCARRILAREGVAAFYKGYVPNMLGIIIPYAGIDLAVYETLKNWLQHYAVNS
ADPGVFVLLACGTMSSTCGQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFG
LYRGLAPNFMKVIPAVSISYVVYENLKITLGVQSR

Important features:**Signal peptide:**

amino acids 1-16

Transmembrane domain:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation site.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

208/330

FIGURE 208

MASLGQILFWSIIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDI
KLSDIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTD
AGTYKCYIIITSKGKGNNANLEYKTGAFSMPEVNVDYNASSETLRCEAPRWFPQPTVVWASQVD
QGANFSEVSNTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTESIARR
SHLQLLNSKASLCVSSFFAISWALLPLSPYLMLK

209/330

FIGURE 209

[illegible]

210/330

FIGURE 210

MAASLGQVLALVLVAALWGGTQPLLKRASAGLQRVHEPTWAQQLQEMKTLFLNTEYLMFPL
LNQCGSLLYYLTLASTDLTLAVPICNSLAIIFTLIVGKALGEDIGGKRKLDYCECGTQLCGS
RHTCVSSFPEPISPEWVRTRPFPILPFPLQLFCFLVAIRVPFPWTVWRKTEAGVWD

211/330

FIGURE 211

CTTCTGTAGGACAGTCACCAGGCCAGATCCAGAAGCCTCTCTAGGCTCCAGCTTTCTCTGTG
GAAGATGACAGCAATTATAGCAGGACCCTGCCAGGCTGTGCGAAAAGATTCCGCAATAAACT
TTGCCAGTGGGAAGTACCTAGTGAAACGGCCTAAGATGCCACTTCTTCTCATGTCCCAGGCT
TGAGGCCCTGTGGTCCCCATCCTTGGGAGAAGTCAGCTCCAGCACCATGAAGGGGCATCCTCG
TTGCTGGTATCACTGCAGTGCTTGTTGCAGCTGTAGAATCTCTGAGCTGCGTGCAGTGTAAT
TCATGGGAAAAATCCTGTGTCAACAGCATTGCCTCTGAATGTCCCTCACATGCCAACACCAG
CTGTATCAGCTCCTCAGCCAGCTCCTCTCTAGAGACACCAGTCAGATTATACCAGAATATGT
TCTGCTCAGCGGAGAACTGCAGTGAGGAGACACACATTACAGCCTTCACTGTCCACGTGTCT
GCTGAAGAACACTTTTCATTTTGTAAAGCCAGTGCTGCCAAGGAAAGGAATGCAGCAACACCAG
CGATGCCCTGGACCCTCCCCTGAAGAACGTGTCCAGCAACGCAGAGTGCCCTGCTTGTTATG
AATCTAATGGAACCTCCTGTCTGTTGGGAAGCCCTGGAAATGCTATGAAGAAGAACAGTGTGTC
TTTCTAGTTGCAGAACTTAAGAATGACATTGAGTCTAAGAGTCTCGTGCTGAAAGGCTGTTC
CAACGTCAGTAACGCCACCTGTCAGTTCCTGTCTGGTGAAAACAAGACTCTTGAGGAGTCA
TCTTTCGAAAGTTTGAGTGTGCAAATGTAAACAGCTTAACCCCCACGTCTGCACCAACCACT
TCCCACAACGTGGGCTCCAAAGCTTCCCTCTACCTCTTGGCCCTTGCCAGCCTCCTTCTTCG
GGGACTGCTGCCCCTGAAGGTCCTGGGGCTGCACTTTGCCCAGCACCCCATTTCTGCTTCTCTG
AGGTCCAGAGCACCCCCTGCGGTGCTGACACCCTCTTCCCTGCTCTGCCCCGTTTAACTGC
CCAGTAAGTGGGAGTCACAGGTCTCCAGGCAATGCCGACAGCTGCCTTGTTCTTCATTATTA
AAGCACTGGTTCATTCACTGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

212/330

FIGURE 212

MKGILVAGITAVLVAAVESLSCVQCNSWEKSCVNSIASECPSHANTSCISSSASSSLETPVR
LYQNMFCSAENCSEETHITAFTVHVSAEEHFHFVSQCCQGKECSNTSDALDPPLKNVSSNAE
CPACYESNGTSCRGKPWKCYEEEQCVFLVAELKNDIESKSLVLKGCSNVSNATCQFLSGENK
TLGGVIFRKFEKANVNSLTPTSAPTTSHNVGSKASLYLLALASLLLRGLLP

213/330

FIGURE 213

GGCCTCGGTTCAAACGACCCGGTGGGTCTACAGCGGAAGGGAGGGAGCGAAGGTAGGAGGCA
GGGCTTGCCTCACTGGCCACCCTCCCAACCCCAAGAGCCCAGCCCC**ATG**GTCCCCGCCGCCG
GCGCGCTGCTGTGGGTCTGCTGCTGAATCTGGGTCCCCGGGCGGCGGGGGCCCAAGGCCTG
ACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGCTTTGGGGGCCCCATGACCCGCAG
CTACCGGAGCACCGCCCGGACTGGTCTTCCCCGGAAGACAAGGATAATCCTAGAGGACGAGA
ATGATGCCATGGCCGACGCCGACCGCCTGGCTGGACCAGCGGCTGCCGAGCTCTTGGCCGCC
ACGGTGTCCACCGGCTTTAGCCGGTCTGCCGCATTAACGAGGAGGATGGGTCTTCAGAAGA
GGGGGTGTGATTAATGCCGGAAGGATAGCACCAGCAGAGAGCTTCCCAGTGCGACTCCCA
ATACAGCGGGGAGTTCCAGCACGAGGTTTATAGCCAATAGTCAGGAGCCTGAAATCAGGCTG
ACTTCAAGCCTGCCGCGCTCCCCCGGAGGTCTACTGAGGACCTGCCAGGCTCGCAGGCCAC
CCTGAGCCAGTGGTCCACACCTGGGTCTACCCCGAGCCGGTGGCCGTCACCCTCACCCACAG
CCATGCCATCTCCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCGTGGCACTGCCAC
TGCAAGTCGGGCACCATGAGCCGGAGCCGGTCTGGGAAGCTGCACGGCCTTCCGGGCGCCT
TCGAGTTGGGGCGCTGAGCCAGCTCCGCACGGAGCACAAGCCTTGACCTATCAACAATGTC
CCTGCAACCGACTTCGGGAAGAGTGCCCCCTGGACACAAGTCTCTGTACTGACACCAACTGT
GCCTCTCAGAGCACCAACAGTACCAGGACCACCACTACCCCCTTCCCCACCATCCACCTCAG
AAGCAGTCCCAGCCTGCCACCCGCCAGCCCCTGCCAGCCCTGGCTTTTTTGGAACGGGTCA
GGATTGGCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTACAGAGATGCAACCAATA
GACAGAAACCAGAGG**TAA**TGGCCACTTCATCCACATGAGGAGATGTCAGTATCTCAACCTCT
CTTGCCCTTTCAATCCTAGCACCCACTAGATATTTTTAGTACAGAAAAACAAAACCTGGAAAA
CACAA

214/330

FIGURE 214

MVPAAGALLWVLLLNLGPRAGAQGLTQTPTEMQRVSLRFGGPMTRSYRSTARTGLPRKTRI
ILEDENDAMADADRLAGPAAAEELLAATVSTGFSRSSAINEEDGSSEEGVVINAGKDSTSREL
PSATPNTAGSSSTRFIANSQEPEIRLTSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWP
SPSPTAMPSPEDLRLVLMPWGPWHCHCKSGTMSRSRSGKLHGLSGRLRVGALSQLRTEHKPC
TYQQPCPNRLREECPLDTSLCDTNCASQSTTSTRTTTTTFPTIHLRSSPSLPPASPCPALA
FWKRVRIGLEDIWNSLSSVFTEMQPIDRNQR

215/330

FIGURE 215

CCCGGGTCGACCCACGCGTCCGGGGAGAAAGGATGGCCGGCCTGGCGGCGCGGTTGGTCCTGCTAGCTGGGGCA
GCGGCGCTGGCGAGCGGCTCCCAGGGCGACCGTGAGCCGGTGTACCGCGACTGCGTACTGCAGTGCGAAGAGCA
GAACTGCTCTGGGGGCGCTCTGAATCACTTCCGCTCCCGCCAGCCAATCTACATGAGTCTAGCAGGCTGGACCT
GTCGGGACGACTGTAAGTATGAGTGTATGTGGGTCACCGTTGGGCTCTACCTCCAGGAAGGTCACAAAGTGCCCT
CAGTTCATGGCAAGTGCCCTTCTCCCGGTTCTGTCTTTCAAGAGCCGGCATCGGCGCTGGCCCTCGTTTCT
CAATGGCCTGGCCAGCCTGGTGATGCTCTGCCGCTACCGCACCTTCGTGCCAGCCTCCTCCCCATGTACCACA
CCTGTGTGGCCTTCGCCTGGGTGTCCCTCAATGCATGGTTCTGGTCCACAGTCTTCCACACCAGGGACACTGAC
CTCACAGAGAAAATGGACTACTTCTGTGCCTCCACTGTCATCCTACACTCAATCTACCTGTGCTGCGTCAGGAC
CGTGGGGCTGCAGCACCCAGCTGTGGTCAGTGCCTTCCGGGCTCTCCTGCTGCTCATGCTGACCGTGACAGTCT
CCTACCTGAGCCTCATCCGCTTCGACTATGGCTACAACCTGGTGGCCAACGTGGCTATTGGCCTGGTCAACGTG
GTGTGGTGGCTGGCCTGGTGCCTGTGGAACCAGCGGCGGCTGCCTCACGTGCGCAAGTGCGTGGTGGTGGTCTT
GCTGCTGCAGGGGCTGTCCCTGCTCGAGCTGCTTGACTTCCCACCGCTCTTCTGGGTCTGGATGCCCATGCCA
TCTGGCACATCAGCACCATCCCTGTCCACGTCTCTTTTTTCAGCTTTCTGGAAGATGACAGCCTGTACCTGCTG
AAGGAATCAGAGGACAAGTTCAAGCTGGACTGAAGACCTTGGAGCGAGTCTGCCCCAGTGGGGATCCTGCCCC
GCCCTGCTGGCCTCCCTTCTCCCCCTCAACCTTGAGATGATTTTCTCTTTTCAACTTCTTGAACCTGGACATGA
AGGATGTGGGCCCAGAATCATGTGGCCAGCCACCCCTGTTGGCCCTCACCAGCCTTGAGTCTGTTCTAGGG
AAGGCTCCAGCATCTGGGACTCGAGAGTGGGCAGCCCTCTACCTCCTGGAGCTGAAGTGGGTTGGAAGTGA
GTGTGTTCTTAGCTCTACCGGGAGGACAGCTGCCTGTTTCTCCCCACCAGCCTCCTCCCCACATCCCCAGCTG
CCTGGCTGGGTCTGAAGCCCTCTGTCTACCTGGGAGACCAGGGACCACAGGCCTTAGGGATACAGGGGGTCCC
CTTCTGTACCACCCCCACCCCTCCTCCAGGACACCACTAGGTGGTGTGCTGGATGCTTGTCTTTGGCCAGCCAA
GGTTCACGGCGATTCTCCCCATGGGATCTTGAGGGACCAAGCTGCTGGGATTGGGAAGGAGTTTACCCTGACC
GTTGCCCTAGCCAGGTTCCAGGAGGCCTCACCATACTCCCTTTCAGGGCCAGGGCTCCAGCAAGCCAGGGCA
AGGATCCTGTGCTGCTGTCTGGTTGAGAGCCTGCCACCGTGTGTGCGGAGTGTGGGCCAGGCTGAGTGCATAGG
TGACAGGGCCGTGAGCATGGGCCTGGGTGTGTGTGAGCTCAGGCCTAGGTGCGCAGTGTGGAGACGGGTGTTGT
CGGGGAAGAGGTGTGGCTTCAAAGTGTGTGTGTGCAGGGGGTGGGTGTGTTAGCGTGGGTTAGGGGAACGTGTG
TGCGCGTGCTGGTGGGCATGTGAGATGAGTGACTGCCGGTGAATGTGTCCACAGTTGAGAGGTTGGAGCAGGAT
GAGGGAATCCTGTACCATCAATAATACTTGTGGAGCGCCAGCTCTGCCCCAAGACGCCACCTGGGCGGACAGC
CAGGAGCTCTCCATGGCCAGGCTGCCTGTGTGCATGTTCCCTGTCTGGTGGCCCTTGGCCGCTCCTGCAAAC
CTCACAGGGTCCCCACACAACAGTGCCCTCCAGAAGCAGCCCTCGGAGGCAGAGGAAGGAAAATGGGGATGGC
TGGGGCTCTCTCCATCCTCCTTTTCTCCTTGCCCTTCGCATGGCTGGCCTTCCCCTCCAAAACCTCCATTCCCCT
GCTGCCAGCCCCTTTGCCATAGCCTGATTTTGGGGAGGAGGAAGGGGCGATTTGAGGGAGAAGGGGAGAAAGCT
TATGGCTGGGTCTGGTTTCTTCCCTTCCCAGAGGGTCTTACTGTTCCAGGGTGGCCCCAGGGCAGGCAGGGGCC
ACACTATGCCCTGTGCCCTGGTAAAGGTGACCCCTGCCATTTACCAGCAGCCCTGGCATGTTCTTCCCCACAGG
AATAGAATGGAGGGAGCTCCAGAACTTTCCATCCCAAAGGCAGTCTCCGTGGTTGAAGCAGACTGGATTTTGTG
CTCTGCCCCTGACCCCTTGTCCCTCTTTGAGGGAGGGGAGCTATGCTAGGACTCCAACCTCAGGGACTCGGGTG
GCCTGCGCTAGCTTCTTTTGATACTGAAAACTTTAAAGGTGGGAGGGTGGCAAGGGATGTGCTTAATAAATCAA
TTCCAAGCCTCAAAAAAAAAAAAAAAAAA

216/330

FIGURE 216

MAGLAARLVLLAGAAALASGSQGDREPVYRDCVLQCEEQNCSGGALNHFRSRQPIYMFLAGW
TCRDDCKYECMWVTVGLYLQEGHKVPQFHGKWPFSSRFLFFQEPASAVASFLNGLASLVMLCR
YRTFVPASSPMYHTCVAFAWVSLNAWFWSTVFHTRDSDLTEKMDYFCASTVILHSIYLCCVR
TVGLQHPAVVSAFRALLLLMLTVHVSYSLSLIRFDYGYNLVANVAIGLVNVVWWLAWCLWNQR
RLPHVRKCVVVVLLQLGLSLLELLDFPPLFWVLDAAHAIWHISTIPVHVLFFSFLEDDSLYLL
KESEDKFKLD

Important features:**Signal peptide:**

amino acids 1-20

Transmembrane domains:

amino acids 105-123, 138-156, 169-185, 193-209, 221-240, 256-272

N-glycosylation site.

amino acids 40-44

N-myristoylation site.

amino acids 43-49

CUB domain proteins profile.

amino acids 285-302

Amiloride-sensitive sodium channels proteins.

amino acids 162-186

217/330

FIGURE 217

[illegible]

218/330

FIGURE 218

MAPQSLPSSRMAPLGMLLGLLMAACFTFCLSHQNLKEFALTNPEKSSTKETERKETKAEEEL
DAEVLEVVFHPTHEWQALQPGQAVPAGSHVRLNLQTGEREAKLQYEDKFRNNLKGKRLDINTN
TYTSQDLKSALAKFKEGAEMESSKEDKARQAEVKRLFRPIEELKKDFDELNVVIETDMQIMV
RLINKFNSSSSSLEEKIAALFDLEYVHQMDNAQDLLSFGGLQVVINGLNSTEPLVKEYAAF
VLGAAFSSSNPKVQVEAIEGGALQKLLVILATEQPLTAKKKVLFALCSLLRHFPYAQRQFLKL
GGLQVLRRTLQVEKGTEVLAVRVVTLVLDLVTEKMFEEEEAELTQEMSPEKLQQYRQVHLLPG
LWEQGWCEITAHLLALPEHDAREKVLQTLGVLLTTCRDRYRQDPQLGRTLASLQAEYQVLAS
LELQDGEDEGYFQELLGSVNSLLKELR

Important features:**Signal peptide:**

amino acids 1-29

Hypothetical YJL126w/YLR351c/yhcX family protein.

amino acids 364-373

N-glycosylation site.

amino acids 193-197, 236-240

N-myristoylation site.

amino acids 15-21, 19-25, 234-240, 251-257, 402-408, 451-457

Homologous region SLS1 protein.

amino acids 68-340

219/330

FIGURE 219

TTCGGCTTCCGTAGAGGAAGTGGCGCGGACCTTCATTTGGGGTTTCGGTTCCCCCCTTCCC
CTTCCCCGGGGTCTGGGGGTGACATTGCACCGCGCCCCCTCGTGGGGTTCGCGTTGCCACCCCA
CGCGGACTCCCCAGCTGGCGCGCCCCCTCCCATTTGCCTGTCTGGTCAGGCCCCCACC
TTCCCACCTGACCAGCC**ATG**GGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTCGCGTTTCGGC
CCGGCCTTCGCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGT
CGCAGGGGCATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGG
TCCATGTGACCGACCGGTGAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCT
GTCTCTGTCTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGA
TGAAGGGTTAGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCT
ATGTTTCTGGTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCT
GATGCACTTGGGCCAGGTGTGGTTGGGATCCATGGAGACTCACCTATTACTTCTGACTTC
AGCCTTTCTGACAGCAGCCATTATCCTGCTCCATACCTTTTGGGGAGTTGTGTTCTTTGATG
CCTGTGAGAGGAGACGGTACTGGGCTTTGGGCCTGGTGGTTGGGAGTCACCTACTGACATCG
GGACTGACATTCTGAACCCCTGGTATGAGGCCAGCCTGCTGCCCATCTATGCAGTCACTGT
TTCCATGGGGCTCTGGGCCTTCATCACAGCTGGAGGGTCCCTCCGAAGTATTCAGCGCAGCC
TCTTGTGTAAGGACT**TGA**CTACCTGGACTGATCGCCTGACAGATCCCACCTGCCTGTCCACTG
CCCATGACTGAGCCCAGCCCCAGCCCCGGGTCCATTGCCACATTCTCTGTCTCCTTCTCGTC
GGTCTACCCCACTACCTCCAGGGTTTTGCTTTGTCTTTTGTGACCGTTAGTCTCTAAGCTT
TACCAGGAGCAGCCTGGGTTTCCAGCCAGTCAGTGACTGGTGGGTTTGAATCTGCACTTATCCC
CACCACCTGGGGACCCCTTGTGTGTCCAGGACTCCCCCTGTGTGCTGCTCTGCTCTCAC
CCTGCCCAAGACTCACCTCCCTTCCCCTCTGCAGGCCGACGGCAGGAGGACAGTCGGGTGAT
GGTGTATTCTGCCCTGCGCATCCCACCCGAGGACTGAGGGAACCTAGGGGGGACCCCTGGGC
CTGGGGTGCCCTCCTGATGTCCTCGCCCTGTATTTCTCCATCTCCAGTTCTGGACAGTGCAG
GTTGCCAAGAAAAGGGACCTAGTTTAGCCATTGCCCTGGAGATGAAATTAATGGAGGCTCAA
GGATAGATGAGCTCTGAGTTTCTCAGTACTCCCTCAAGACTGGACATCTTGGTCTTTTTCTC
AGGCCTGAGGGGGAACCATTTTTGGTGTGATAAATAACCTAAACTGCCTTTTTTTCTTTTTT
GAGGTGGGGGGAGGGAGGAGGTATATTGGAACCTCTTAACCTCCTTGGGCTATATTTCTC
TCCTCGAGTTGCTCCTCATGGCTGGGCTCATTTTCGGTCCCTTTCTCCTTGGTCCCAGACCTT
GGGGGAAAGGAAGGAAGTGCATGTTTGGGAACCTGGCATTACTGGAACATAATGGTTTTAACCT
CCTTAACCACAGCATCCCTCCTCTCCCCAAGGTGAAGTGGAGGGTGCTGTGGTGAGCTGGC
CACTCCAGAGCTGCAGTGCCACTGGAGGAGTCAGACTACCATGACATCGTAGGGAAGGAGGG
GAGATTTTTTTGTAGTTTTTAATTGGGGTGTGGGAGGGGCGGGGAGGTTTTCTATAAACTGT
ATCATTTTCTGCTGAGGGTGGAGTGTCCCATCCTTTTAATCAAGGTGATTGTGATTTTGACT
AATAAAAAAGAATTTGTAAAAA
AAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

220/330

FIGURE 220

MGA AVFFGCTFVAFGPAFALFLITVAGDPLRVIIILVAGAFFWLVSLLLASVVWFILVHVTDR
SDARLQYGLLIFGA AVSVLLQEVFRFAYYKLLKKADEGLASLSEDGRSPISIRQMAYVSGLS
FGIISGVF SVINILADALGPGVVG I HGDS PYYFLTSAFLTAAIILLHTFWGVVFFDACERRR
YWALGLVVGSHLLTSGLTFLNPWYEASLLPIYAVTVSMGLWAFITAGGSLRSIQRSLLCKD

221/330

FIGURE 221

AAGCTGGTTTAAGGAAGCAGAGGAGGGTTAGATTCGTTGAGTGAGGACGGAAGATCAACCCA
TTTCCATTCCGCCAGATGGCCTATGTTTCTGGTCTCTCCCTTCGGNATCATCAGTGGTGTNT
TNTCTGTTATCAATATTTTGGCTGATGCANTTGGGCCAGGTGTGGTTGGGATCCATGGAGAC
TCACCCTATTANTTCCTGANTTCAGCCTTTNTGACAGCAGCCATTATCCTGCTC

222/330

FIGURE 222

GACCGACCGTTCAGATGCCCCGGTTCAGTACGGCTTCCTGATTTTTGGTGCTGCTGTNTCTG
TCCTTCTACAGGAGGTGTTCCGCTTTGCCTANTACAAGCTGCTTAAGAAGGCAGATGAGGGG
TTAGCATNGCTGAGTGAGGACGGAAGATCACCCATTTCCATCCGCCAGATGGCCTATGTTTN
TGGTNTTTCCTTCGGTATCATCAGTGGTGTNTNTCTGTTATCAATATTTGGNTGATGCAN
TTGGGCCAGGTGTGGTTGGGATCCATGGAGANTCACCTATTAATTCCTGAATTCAGCCTTT
NTGACAGCAGCCATTATCCTGNTCCATACCTTTTGGGGAGTTGTGTTTTTTGATGCCTGTGA
GAGGAG

223/330

FIGURE 223

NGTTGGAGAAGTGGCGCGGACNTTCATTTGGGGTTTCGGTTTCCCCCCTTTCCCTTTCCCCG
GGGTCTGGGGTGACATTGCACGGGCCCCCTCGTGGGGTCGCGTTGCCACCCACGCGGACTCC
CCAGNTGGNGCGCCCTTCCCATTTGCCTGTCCTGGTCAGGCCCCACCCCCCTTCCACNTG
ACCAGCCATGGGGGCTGCGGTGTTTTTCGGCTGCACTTTCGTCGCGTTCGGCCCGGCCTTCG
CGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGCA
TTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGAC
CGACCGGTCAGATGCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTCC
TTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTTA
GCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGG
TCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTGGCTGATGCACTTG
GGCCAGGTGTGGTTGGGATCCATGGAGACTCACCC

224/330

FIGURE 224

GTAAAAGAAAGTGGCCGGACCTTCATTGGGGTTTCGGTTCCCCCCTTCCCNTTCCCCGGGG
TCTGGGGGTGACATTGCACCGCGCCCNCTCGTGGGGTCGCGTTGCCACCCCACGCGGACTCCC
CAGNTGGCGCGCCCCCTCCCATTTGCCTGTCCTGGTCAGGCCCCCACCCTTCCCACCTGA
CCAGCCATGGGGGCTGCGGTGTTTTTCGGGCTGCACTTTCGTCGCGTTCGGGCCCCGGCCTTC
GCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTATCATCCTGGTCGCAGGGGC
ATTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCCATGTGA
CCGACCGGTCAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTGGTGCTGCTGTCTCTGTC
CTTCTACAGGAGGTGTTCCGCTTTGCCTACTACAAGCTGCTTAAGAAGGCAGATGAGGGGTT
AGCATCGCTGAGTGAGGACGGAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTG
GTCTCTCCTTCGGTATCATCAGTGGTGTCTTCTCTGTTATCAATATTTTGGCTGATGCACTT
GGGCCAGGTGTGGTTGGGATCCATGGAGAC

225/330

FIGURE 225

GCCCCAGGGAGCAGTGGGTGGTTATAACTCAGGCCCCGGTGCCCAGAGCCCAGGAGGAGGCAG
TGGCCAGGAAGGCACAGGCCTGAGAAGTCTGCGGCTGAGCTGGGAGCAAATCCCCACCCCC
TACCTGGGGGACAGGGCAAGTGAGACCTGGTGAGGGTGGCTCAGCAGGCAGGGAAGGAGAGG
TGTCTGTGCGTCCTGCACCCACATCTTTCTCTGTCCCCTCCTTGCCCTGTCTGGAGGCTGCT
AGACTCCTATCTTCTGAATTCTATAGTGCCTGGGTCTCAGCGCAGTGCCGATGGTGGCCCCGT
CCTTGTGGTTCCTCTCTACCTGGGGAAATAAGGTGCAGCGGCC**ATG**GCTACAGCAAGACCCC
CCTGGATGTGGGTGCTCTGTGCTCTGATCACAGCCTTGCTTCTGGGGGTACAGAGCATGTT
CTCGCCAACAATGATGTTTCCTGTGACCACCCCTCTAACACCGTGCCCTCTGGGAGCAACCA
GGACCTGGGAGCTGGGGCCGGGAAGACGCCCGGTCGGATGACAGCAGCAGCCGCATCATCA
ATGGATCCGACTGCGATATGCACACCCAGCCGTGGCAGGCCGCGCTGTTGCTAAGGCCCAAC
CAGCTCTACTGCGGGGCGGTGTTGGTGCATCCACAGTGGCTGCTCACGGCCGCCCACTGCAG
GAAGAAAGTTTTAGAGTCCGTCTCGGCCACTACTCCCTGTCACCAGTTTATGAATCTGGGC
AGCAGATGTTCCAGGGGGTCAAATCCATCCCCACCCTGGCTACTCCCACCCTGGCCACTCT
AACGACCTCATGCTCATCAAACGAAGAAGTTCGTCCCACTAAAGATGTCAGACCCAT
CAACGTCTCCTCTCATTGTCCCTCTGCTGGGACAAAGTGCTTGGTGTCTGGCTGGGGGACAA
CCAAGAGCCCCCAAGTGCACTTCCCTAAGGTCCTCCAGTGCTTGAATATCAGCGTGCTAAGT
CAGAAAAGGTGCGAGGATGCTTACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGA
CAAAGCAGGTAGAGACTCCTGCCAGGGTGATTCTGGGGGGCCTGTGGTCTGCAATGGCTCCC
TGCAGGGACTCGTGTCTTGGGGAGATTACCCCTTGCGCCGGCCCAACAGACCGGGTGTCTAC
ACGAACCTCTGCAAGTTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC**TGA**GTCAT
CCCAGGACTCAGCACACCGGCATCCCCACCTGCTGCAGGGACAGCCCTGACACTCCTTTTCAG
ACCCTCATTCTTCCCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCCTGACCCCATGTCT
CCTGGACTCAGGGTCTGCTTCCCCACATTGGGCTGACCGTGTCTCTCTAGTTGAACCCTGG
GAACAATTTCCAAAACGTGTCAGGGCGGGGGTTGCGTCTCAATCTCCCTGGGGCACTTTTCAT
CCTCAAGCTCAGGGGCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCCAGAAATAAA
CTGAGAAGTGGAATAAAAAA

226/330

FIGURE 226

MATARPPWMWVLCALITALLLGVTEHVLANNNDVSCDHPSNTVPSGSNQDLGAGAGEDARSDD
SSSRIINGSDCDMHTQPWQAALLLRPNQLYCGAVLVHPQWLLTAAHCRKKVFRVRLGHYSLS
PVYESGQQMFQGVKSI PHPGYSHPGHSNDLMLIKLNRRIRPTKDVRPINVSSHCP SAGTKCL
VSGWGTTKSPQVHF PKVLQCLNISVLSQKRCEDAYPRQIDDTMFCAGDKAGRDSCQGDSGGP
VVCNGSLQGLVSWGDYPCARPNRPGVYTNLCKFTKWIQETIQANS

227/330

FIGURE 227

ATGGTCAACGACCGGTGGAAGACCATGGGCGGGCGCTGCCCAACTTGAGGACCGGCCGCGCGA
CAAGCCCGCAGCGGCCGAGCTGCGGGCTACGTGCTGTGCACCGTGCTGCTGGCCCTGGCTGTGC
TGCTGGCTGTAGCTGTACCGGTGCCGTGCTCTTCCTGAACCACGCCCACGCGCCGGGCACG
GCGCCCCACCTGTCGTACGCACTGGGGCTGCCAGCGCCAACAGCGCCCTGGTCACTGTGGA
AAGGGCGGACAGCTCGCACCTCAGCATCCTCATTGACCCGCGCTGCCCCGACCTCACCGACA
GCTTCGCACGCCTGGAGAGCGCCCAGGCCCTCGGTGCTGCAGGCGCTGACAGAGCACCAGGCC
CAGCCACGGCTGGTGGGCGACCAGGAGCAGGAGCTGCTGGACACGCTGGCCGACCAGCTGCC
CCGGCTGCTGGCCCGAGCCTCAGAGCTGCAGACGGAGTGCATGGGGCTGCGGAAGGGGCATG
GCACGCTGGGCCAGGGCCTCAGCGCCCTGCAGAGTGAGCAGGGCCGCTCATCCAGCTTCTC
TCTGAGAGCCAGGGCCACATGGCTCACCTGGTGAACCTCCGTACGCGACATCCTGGATGCCCT
GCAGAGGGACCGGGGGCTGGGCCGGCCCCGCAACAAGGCCGACCTTCAGAGAGCGCCTGCCC
GGGGAACCCGGCCCCGGGGCTGTGCCACTGGCTCCCGGCCCGAGACTGTCTGGACGTCTCT
CTAAGCGGACAGCAGGACGATGGCGTCTACTCTGTCTTCCCACCCACTACCCGGCCGGCTT
CCAGGTGTACTGTGACATGCGCAGGACGCGCGCGCTGGACGGTGTTCAGCGCCGGGAGG
ACGGCTCCGTGAACCTTCTCCGGGGCTGGGACGCGTACCAGAGACGGCTTTGGCAGGCTCACC
GGGGAGCACTGGCTAGGGCTCAAGAGGATCCACGCCCTGACCACACAGGCTGCCTACGAGCT
GCACGTGGACCTGGAGGACTTTGAGAATGGCAGGCGCTATGCCCGCTACGGGAGCTTCGGCG
TGGGCTTGTCTCCGTGGACCCTGAGGAAGACGGGTACCCGCTCACCGTGGCTGACTATTCC
GGCACTCAGGCGACTCCCTCCTGAAGCACAGCGGCATGAGGTTACCAACCAAGGACCCGTGA
CAGCGACCATTCAGAGAACAACCTGTGCCGCCCTTCTACCGCGGTGCCTGGTGGTACCGCAACT
GCCACACGTCCAACCTCAATGGGCAGTACCTGCGCGGTGCGCACGCCTCCTATGCCGACGGC
GTGGAGTGGTCTCTGACCCGGCTGGCAGTACTCACTCAAGTTCTCTGAGATGAAGATCCG
GCCGGTCCGGGAGGACCGCT**TAG**ACTGGTGCACCTTGTCTTGGCCCTGCTGGTCCCTGTGCG
CCCATCCCCGACCCACCTCACTCTTTCGTGAATGTTCTCCACCCACCTGTGCCCTGGCGGAC
CCACTCTCCAGTAGGGAGGGGCGGGCCATCCCTGACACGAAGCTCCCTGGGCCGGTGAAGT
CACACATCGCCTTCTCGCCGTCCCCACCCCTCCATTTGGCAGCTCACTGATCTCTTGCCTC
TGCTGATGGGGGCTGGCAAACCTTGACGACCCCAACTCCTGCCTGCCCCACTGTGACTCCGG
TGCTGTTTGGCGTCCCCCTGGCCAGGATGGTGGAGTCTGCCCCAGGCACCCCTCTGCCCTGCCC
GGCCAAATACCCGCGATTATGGGGACAGAGAGCGGGGCGAGACAGACCCCTGGAGTCTCT
CTAGCAGATCGTGGGAATGTCAAGGTCTCTCTGAGGTCAAGTCTGAGGCCAGTATCCTCCAG
CCCTCCCAATGCCAACCCCCACCCGTTTCCCTGGTGCCAGAGAACCCACCTCTCCCCCAA
GGGCCTCAGCCTGGCTGTGGGCTGGGTGGCCCCATCCTACCAGGCCCTGAGGTCAGGATGGG
GAGCTGCTGCCTTTGGGGACCCACGCTCCAAGGTGAGACCAGTTCCTTGGAGGCCACCCAC
CCTGTGCCCCGGCAGGCCTGGGGTCTGCAGTCTCTTACCTGCTGTGCCCACCTGCTCTCTG
TCTCAAATGAGGCCCAACCCATCCCCACCCAGCTCCCGGCCGTCTCTTACCTGGGGCAGC
CGGGGCTGCCATCCCATTTCTCCTGCCTCTGGAAGGTGGGTGGGGCCCTGCACCGTGGGGCT
GGACTGCGCTAATGGGAAGCTCTTGGTTTTCTGGGCTGGGGCCTAGGCAGGGCTGGGATGAG
GCTTGTACAACCCCCACCACCAATTTCCCAGGGACTCCAGGGTCTGAGGCCTCCCAGGAGG
GCCTTGGGGGTGATGACCCCTTCCCTGAGGTGGCTGTCTCCATGAGGAGGCCAACCCTTGCC
ATTGACCGTGGCCACCTGGACCCAGGCCAGGCCCGGCCGCGAGTGGTCAAGGGACAGGGA
CCACCTCACCGGGCAAATGGGGTGGGGGGGACTGGGGCACCAGACCAGGCACCACTGGACA
CTTTCTTGTGAATCCTCCCAACACCCAGCACGCTGTCATCCCCACTCCTTGTGTGCACACA
TGCAGAGGTGAGACCCGCGAGGCTCCCAGGACCAGCACCAAGGGCAGGGCTGGAGCCGGG
TCCTCAGCTGTCTGCTCAGCAGCCCTGGACCCGCGTGCCTTACGTACGGCCAGATGCAGGG
CGGCTTTTCCAAGGCCTCCTGATGGGGGGCTCCGAAAGGGCTGGAGTCAGCCTTGGGGAGCT
GCCTAGCAGCCTCTCCTCGGGCAGGAGGGGAGGTGGCTTCTCCTCAAAGGACACCCGATGGCA
GGTGCCTAGGGGGTGTGGGGTTCCGTTCTCCCTTCCCTCCCACTGAAGTTTGTGCTTAAAA
AACAAATAAATTTGACTTGGCACCACTGGGGGTTGGTGGGAGAGGCCGTGTGACCTGGCTCTC
TGTCCAGTGCCACCAGGTCATCCACATGCGCAG

228/330

FIGURE 228

MVNDRWKTMGGAAQLEDPRDPKQRPSCGYVLCTVLLALAVLLAVAVTGAVLFLNHAHAPGT
APPPVVSTGAASANSALVTVERADSSHLSILIDPRCPDLTDSFARLESAQASVLQALTEHQA
QPRLVGDQEQELLDLADQLPRLLARASELQTECMGLRKGHGTLGQGLSALQSEQGRLIQLL
SESQGHMAHLVNSVSDILDALQDRGLGRPRNKADLQRAPARGTRPRGCATGSRPRDCLDVL
LSGQQDDGVYSVFPTHYPAGFQVYCDMRTDGGGWTVFQRREDGSVNFFRGWDAYRDGFGRLT
GEHWLGLKRIHALTTQAAYELHVDLEDFENGTAAYARYGSFGVGLFSVDPEEDGYPLTVADYS
GTAGDSLLKHSGMRFTTKDRSDHSENNCAAFYRGAWWYRNCHTSNLNGQYLRGAHASADG
VEWSSWTGWQYSLKFSEM KIRPVREDR

229/330

FIGURE 229

GCAGTCAGAGACTTCCCCTGCCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCT
TGCTTCCTGAACTAGCTCACAGTAGCCCGGCGGCCAGGGCAATCCGACCACATTTCACTCT
CACCGCTGTAGGAATCCAG**ATG**CAGGCCAAGTACAGCAGCACGAGGGACATGCTGGATGATG
ATGGGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACCTCGGCATCCAGAGCCC
CGGCGCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGAC
TTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTTCAGTACTACC
AGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATACGTCC
CAAGAGTTGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGC
TGAAAAACTCTGTCTGTAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGTACAG
AACAAATGGAAATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAG
GACTGTAAATATTTCTGCCTTAGTGAAAACCTCTACCATGCTGAAGATAAACAAACAAGAAGA
CCTGGAATTTGCCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTT
TGCGCCCTGACAGTGGCAAGGCCTGGCTGTGGATGGATGGAACCCCTTTCACTTCTGAACTG
TTCCATATTATAATAGATGTCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGG
GATGATCTTCTCAAAGGACTGCAAAGAATTGAAGCGTTGTGTCTGTGAGAGAAGGGCAGGAA
TGGTGAAGCCAGAGAGCCTCCATGTCCCCCTGAAACATTAGGCGAAGGTGACT**TGA**ATTCGCC
CTCTGCAACTACAAATAGCAGAGTGAGCCAGGCGGTGCCAAAGCAAGGGCTAGTTGAGACAT
TGGGAAATGGAACATAATCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCTCGTG
TTTCCTGTTTCAGGATCACCAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACA
AGAAGTCTTATTTACATGCCACCAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTG
GCTTAGAGATAACTTTTAGCTCTCTTTCTTCTCAATGTCTAATATCACCTCCCTGTTTTTCAT
GTCTTCCTTACACTTGGTGGAATAAGAACTTTTTGAAGTAGAGGAAATACATTGAGGTAAC
ATCCTTTTCTCTGACAGTCAAGTAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTACC
AGCAAATACACAAGGAATTCTTTTTGTTTGTTCAGTTCATACTAGTCCCTTCCCAATCCAT
CAGTAAAGACCCCATCTGCCTTGTCCATGCCGTTTCCCAACAGGGATGTCACTTGATATGAG
AATCTCAAATCTCAATGCCTTATAAGCATTCCCTTCCTGTGTCCATTAAGACTCTGATAATTG
TCTCCCCTCCATAGGAATTTCTCCCAGGAAAGAAATATATCCCCATCTCCGTTTCATATCAG
AACTACCGTCCCCGATATTCCCTTCAGAGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCA
TCTGCACCTGTAATAGTTTCAGTTCCTATTTTCTTCCATTGACCCATATTTATACCTTTCAG
GTACTGAAGATTTAATAATAATAAATGTAAATACTGTGAAAAA

230/330

FIGURE 230

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCVL
IGLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKL
LYNKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLE
FAAS
QSYSEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMI
FSKD
CKELKRCVCERRAGMVKPESLHVPPETLGEGD

231/330

FIGURE 231

AATTTTCACCGCTGTAGGAATCCAGATGCAGGCCAAGTACAGCAGCACGAGGGACATGNTGG
ATGATGATGGGACACCACCATGAGCCTGCATTNTCAAGCTTTTGCCACAATTCGGCATCCAG
AGCCCCGGCGCACAGAGCACAGGGNTCCTTTTTCAACGTGGCGACCAGTGGCCCTGACCCTG
CTGACTTTGTGCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTTCAGTA
CTACCAGCTCTCCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATA
CGTCCCAAGAGTTGCAATTTNTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTNTGCAGCAT
GTGGCTGAAAAACTCTGTCGTGAGCTGTATAACAAAGCTGGAGGAACTTTGAAGGAGGGCAA
AGTNTCCTCATNTACTATACACACACCACTTCCC

232/330

FIGURE 232

GCCGAGCGCAAGAACCCTGCGCAGCCCAGAGCAGCTGCTGGAGGGGAATCGAGGCGCGGCTC
CGGGGATTCGGCTCGGGCCGCTGGCTCTGCTCTGCGGGGAGGGAGCGGGCCCGCGCGGG
CCCGAGCCCTCCGGATCCGCCCCCTCCCCGGTCCCGCCCCCTCGGAGACTCCTCTGGCTGCT
CTGGGGGTTTCGCCGGGGCCGGGGACCCGCGGTCCGGGCGCC**ATG**CGGGCATCGCTGCTGCTG
TCGGTGCTGCGGCCCCGAGGGCCCGTGGCCGTGGGCATCTCCCTGGGCTTCACCCTGAGCCT
GCTCAGCGTCACCTGGGTGGAGGAGCCGTGCGGCCCAGGCCCGCCCCAACCTGGAGACTCTG
AGCTGCCGCCGCGCGGCAACACCAACGCGGCGCGCCGCCAACCTCGGTGCAGCCCCGAGCG
GAGCGCGAGAAGCCCCGGGGCCGGCGAAGGCGCCGGGGAGAATTGGGAGCCGCGCGTCTTGCC
CTACCACCCTGCACAGCCCCGGCCAGGCCGCCAAAAAGGCCGTGAGGACCCGCTACATCAGCA
CGGAGCTGGGCATCAGGCAGAGGCTGCTGGTGGCGGTGCTGACCTCTCAGACCACGCTGCCC
ACGCTGGGCGTGGCCGTGAACCGCACGCTGGGGCACCGGCTGGAGCGTGTGGTGTTCCTGAC
GGGCGCACGGGGCCGCCGGGCCCCACCTGGCATGGCAGTGGTGACGCTGGGCGAGGAGCGAC
CCATTGGACACCTGCACCTGGCGCTGCGCCACCTGCTGGAGCAGCACGGCGACGACTTTGAC
TGGTTCTTCTGGTGCCCTGACACCACCTACACCGAGGCGCACGGCCTGGCACGCCTAACTGG
CCACCTCAGCCTGGCCTCCGCCGCCACCTGTACCTGGGCCGGCCCCAGGACTTCATCGGCG
GAGAGCCCAACCCCGGCCGCTACTGCCACGGAGGCTTTGGGGTGCTGCTGTGCGCATGCTG
CTGCAACAACCTGCGCCCCACCTGGAAGGCTGCCGCAACGACATCGTCAGTGCAGGCCCTGA
CGAGTGGCTGGGTGCTGCTGCTGATGCCACCGGGGTGGGCTGCACTGGTGACCACGAGG
GGGTGCACTATAGCCATCTGGAGCTGAGCCCTGGGGAGCCAGTGACGAGGGGGACCCTCAT
TTCCGAAGTGCCCTGACAGCCCACCCTGTGCGTGACCCTGTGCACATGTACCAGCTGCACAA
AGCTTTCGCCCGAGCTGAACTGGAACGCACGTACCAGGAGATCCAGGAGTTACAGTGGGAGA
TCCAGAATACCAGCCATCTGGCCGTTGATGGGGACCGGGCAGCTGCTTGGCCCGTGGGTATT
CCAGCACCATCCCGCCCGCCTCCCGCTTTGAGGTGCTGCGCTGGGACTACTTCACGGAGCA
GCACGCTTTCTCCTGCGCCGATGGCTCACCCCGCTGCCCACTGCGTGGGGCTGACCGGGCTG
ATGTGGCCGATGTTCTGGGGACAGCTCTAGAGGAGCTGAACCGCCGCTACCACCCGGCCTTG
CGGCTCCAGAAGCAGCAGCTGGTGAATGGCTACCCACGCTTTGATCCGGCCCCGGGGTATGGA
ATACACGCTGGACTTGCAGCTGGAGGCACTGACCCCCAGGGAGGCCCGCCGCCCTCACTC
GCCGAGTGACGCTGCTCCGGCCGCTGAGCCGCGTGAGAGATCTTGCCTGTGCCCTATGTCACT
GAGGCCTCACGTCTCACTGTGCTGCTGCCTCTAGCTGCGGCTGAGCGTGACCTGGCCCCCTGG
CTTCTTGAGGCGCTTTGCCACTGCAGCACTGGAGCCTGGTGATGCTGCGGCAGCCCTGACCC
TGCTGCTACTGTATGAGCCGCGCCAGGCCCAGCGCTGGCCCATGCAGATGTCTTCGCACCT
GTCAAGGCCACGTGGCAGAGCTGGAGCGGCGTTTCCCCGGTGCCCGGGTGCCATGGCTCAG
TGTGCAGACAGCCGCACCCCTCACCACTGCGCCTCATGGATCTACTCTCCAAGAAGCACCCGC
TGGACACACTGTTCTGCTGGCCGGGCCAGACACGGTGCTCACGCCTGACTTCCTGAACCGC
TGCCGCATGCATGCCATCTCCGGCTGGCAGGCCTTCTTTCCCATGCATTTCCAAGCCTTCCA
CCCAGGTGTGGCCCCACCACAAGGGCCTGGGCCCCAGAGCTGGGCCGTGACACTGGCCGCT
TTGATCGCCAGGCAGCCAGCGAGGCCTGCTTCTACAACCTCCGACTACGTGGCAGCCCGTG
CGCCTGGCGGCAGCCTCAGAACAAGAAGAGGAGCTGCTGGAGAGCCTGGATGTGTACGAGCT
GTTCTCTCACTTCTCCAGTCTGCATGTGCTGCGGGCGGTGGAGCCGGCGCTGCTGCAGCGCT
ACCGGGCCCCAGACGTGCAGCGCGAGGCTCAGTGAGGACCTGTACCACCGCTGCCTCCAGAGC
GTGCTTGAGGGCCTCGGCTCCCGAACCCAGCTGGCCATGCTACTCTTTGAACAGGAGCAGGG
CAACAGCACCT**TGA**CCCCACCCTGTCCCCGTGGGCCGTGGCATGGCCACACCCCACTT
CTCCCCAAAACCAGAGCCACCTGCCAGCCTCGCTGGGCAGGGCTGGCCGTAGCCAGACCCC
AAGCTGGCCCCACTGGTCCCCCTCTCTGGCTCTGTGGGTCCCTGGGCTCTGGACAAGCACTGGG
GGACGTGCCCCCAGAGCCACCCACTTCTCATCCCAAACCCAGTTTCCCTGCCCCCTGACGCT
GCTGATTGCGGCTGTGGCCTCCACGTATTTATGCAGTACAGTCTGCCTGACGCCAGCCCTGC
CTCTGGGCCCTGGGGCTGTAGAAGAGTTGTTGGGGAAGGAGGAGCTGAGGAGGGG
GCATCTCCCAACTTCTCCCTTTTGGACCCTGCCGAAGCTCCCTGCCTTTAATAAACTGGCCA
AGTGTGGAAAAA

233/330

FIGURE 233

MRASLLLSVLRPAGPVAVGISLGFTLSLLSVTWVEEPCGPGPPQPGDSELPARGNTNAARRP
NSVQPGAEREKPGAGEGAGENWEPRVLPYHPAQPGQAAKKAVRTRYISTELGIRQRLLVAVL
TSQTTLPTLGVAVNRTLGHRLERVVFLTGARGRRAPPGMAVVTLGEERPIGHLHLALRHLL
QHGDDEFDWFFLVPDTTYTEAHGLARLTGHLSLASAAHLYLGRPQDFIGGEPTPGRYCHGGFG
VLLSRMLLQQLRPHLEGCRNDIVSARPDEWLGRCLDATGVGCTGDHEGVHYSHLELSPGEP
VQEGDPHFERSALTAHPVRDPVHMYQLHKAFARAEELERTYQEIQELQWEIQNTSHLAVDGDRA
AAWPVGIPAPSRPASRFVLRWDYFTEQHAFSCADGSPRCPLRGADRADVADVLTALAEELN
RRYHPALRLQKQQLVNGYRRFDPARGMEYTLDLQLEALTPQGGRRPLTRRVQLLRPLSRVEI
LPVPYVTEASRLTVLLPLAAAERDLAPGFLEAFATAALEPGDAAAALTLLLLYEPRQAQRVA
HADVFAPVKAHVAELERRFPGARVPWLSVQTAAPSPLRLMDLLSKKHPLDTLFLLAGPDTVL
TPDFLNRCRMHAISGWQAFFPMHFQAFHPGVAPPQGPPELGRDTGRFDRQAASEACFYNS
DYVAARGRLAAASEQEEELLES LDVYELFLHFSSLHVLRAVEPALLQRYRAQTCSARLSEDL
YHRCLQSVLEGLGSRTQLAMLLFEQE QGNST

234/330

FIGURE 234

GCTCTGGCCGGCCCCGGCGATTGGTCACCGCCCGCTAGGGGACAGCCCTGGCCTCCTCTGAT
TGGCAAGCGCTGGCCACCTCCCCACACCCCTTGCGAACGCTCCCCTAGTGGAGAAAAGGAGT
AGCTATTAGCCAATTTCGGCAGGGCCCGCTTTTTAGAAAGCTTGATTTCCCTTTGAAGATGAAAG
ACTAGCGGAAGCTCTGCCTCTTTCCCCAGTGGGCGAGGGAACTCGGGGCGATTGGCTGGGAA
CTGTATCCACCCAAATGTCACCGATTTCCTTCCTATGCAGGAAATGAGCAGACCCATCAATAA
GAAATTTCTCAGCCTGGCCGAAAATGGTTGGCCCCACGAAGCCACGACAACCTGGAGGCCAAAG
AGGGTTGCTCAACGCCCCGCCTCATTGGAAAACCAAATCAGATCTGGGACCTATATAGCGTG
GCGGAGGCGGGGCGATGATTGTCGCGCTCGCACCCACTGCAGCTGCGCACAGTCGCATTTCT
TTCCCCGCCCCCTGAGACCCTGCAGCACCATCTGTC**ATG**GCGGCTGGGCTGTTTGGTTTGAGC
GCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGCTCCCGGCCGCCCGCGTCCGCTGGGA
ATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCTGTGGCGGGAAAGCGGCCCCCAGAAC
CGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGACGAAAACCTTGTATGAGAAGAACCCA
GACTCCCATGGTTATGACAAGGACCCCGTTTTGGACGTCTGGAACATGCGACTTGTCTTCTT
CTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACCTTTGTGGCCTATCTGCCTGACTACA
GGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCTTGTGAAATACCGAGAGGCCAATGGC
CTTCCCATCATGGAATCCAACCTGCTTCGACCCCAGCAAGATCCAGCTGCCAGAGGATGAG**TG**
ACCAGTTGCTAAGTGGGGCTCAAGAAGCACCGCCTTCCCCACCCCCTGCCTGCCATTCTGAC
CTCTTCTCAGAGCACCTAATTAAAGGGGCTGAAAGTCTGAA

235/330

FIGURE 235

MAAGLFGLSARRLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPPEPTTPWQEDPEPE
DENLYEKNPDSHGYDKDPVLDVWNMRLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAER
LVKYREANGLPIMESNCFDPSKIQLPEDE

236/330

FIGURE 236

GGCGGCTGGGCTGTTTGGTTTGAGCGCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGC
TCCCGGCCGCCCCGCGTCCGCTGGGAATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCT
GTGGCGGGAAAGCGGCCCCCAGAACCGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGA
CGAAAACCTTGTATGAGAAGAACCCAGACTCCCATGGTTATGACAAGGACCCCGTTTTGGACG
TCTGGAACATGCGACTTGTCTTCTTCTTTGGCGTCTCCATCATCCTGGTCCTTGGCAGCACC
TTTGTGGCCTATCTGCCTGACTACAGGATGAAAGAGTGGTCCCGCCGCGAAGCTGAGAGGCT
TGTGAAATACCGAGAGGCCAATGGCCTTCCCATCATGGAATCCAAGTCTTCGACCCCAGCA
AGATCCAG

237/330

FIGURE 237

GCGGCGGCT**ATG**CCGCTTGCTCTGCTCGTCCTGTTGCTCCTGGGGCCCGGCGGGCTGGTGCCT
TGCAGAACCCCCACGCGACAGCCTGCGGGAGGAACTTGTCATCACCCCGCTGCCTTCCGGGG
ACGTAGCCGCCACATTCCAGTTCCGCACGCGCTGGGATTTCGGAGCTTCAGCGGGAAGGAGTG
TCCCATTACAGGCTCTTTCCCAAAGCCCTGGGGCAGCTGATCTCCAAGTATTCTCTACGGGA
GCTGCACCTGTCATTACACAAAGGCTTTTGGAGGACCCGATACTGGGGGCCACCCCTTCCTGC
AGGCCCCATCAGGTGCAGAGCTGTGGGTCTGGTTCCAAGACACTGTCAGTGTGGATAAA
TCTTGGAAGGAGCTCAGTAATGTCCTCTCAGGGATCTTCTGCGCCTCTCTCAACTTCATCGA
CTCCACCAACACAGTCACTCCCACTGCCTCCTTCAAACCCCTGGGTCTGGCCAATGACACTG
ACCACTACTTTCTGCGCTATGCTGTGCTGCCGCGGGAGGTGGTCTGCACCGAAAACCTCACC
CCCTGGAAGAAGCTCTTGCCCTGTAGTTCCAAGGCAGGCCTCTCTGTGCTGCTGAAGGCAGA
TCGCTTGTTCCACACCAGCTACCACTCCCAGGCAGTGCATATCCGCCCTGTTTGCAGAAATG
CACGCTGTACTAGCATCTCCTGGGAGCTGAGGCAGACCCTGTCAGTTGTATTTGATGCCTTC
ATCACGGGGCAGGGAAAGAAAGACTGGTCCCTCTTCCGGATGTTCTCCCGAACCCCTCACGGA
GCCCTGCCCCCTGGCTTCAGAGAGCCGAGTCTATGTGGACATCACCACTACAACCAGGACA
ACGAGACATTAGAGGTGCACCCACCCCGACCACTACATATCAGGACGTCATCCTAGGCACT
CGGAAGACCTATGCCATCTATGACTTGCTTGACACCGCCATGATCAACAACCTCTCGAAACCT
CAACATCCAGCTCAAGTGGAAGAGACCCCCAGAGAATGAGGCCCCCCCCAGTGCCCTTCCTGC
ATGCCCAGCGGTACGTGAGTGGCTATGGGCTGCAGAAGGGGGAGCTGAGCACACTGCTGTAC
AACACCCACCCATACCGGGCCTTCCCGGTGCTGCTGCTGGACACCGTACCCTGGTATCTGCG
GCTGTATGTGCACACCCTCACCATCACCTCCAAGGGCAAGGAGAACAAACCAAGTTACATCC
ACTACCAGCCTGCCCAGGACCGGCTGCAACCCACCTCCTGGAGATGCTGATTCAGCTGCCG
GCCAACTCAGTCACCAAGGTTTCCATCCAGTTTGAGCGGGCGCTGCTGAAGTGGACCGAGTA
CACGCCAGATCCTAACCATGGCTTCTATGTCAGCCCATCTGTCCTCAGCGCCCTTGTGCCCA
GCATGGTAGCAGCCAAGCCAGTGGACTGGGAAGAGAGTCCCCTCTTCAACAGCCTGTTCCCA
GTCTCTGATGGCTCTAACTACTTTGTGCGGCTCTACACGGAGCCGCTGCTGGTGAACCTGCC
GACACCGGACTTCAGCATGCCCTACAACGTGATCTGCCTCACGTGCACTGTGGTGGCCGTGT
GCTACGGCTCCTTCTACAATCTCCTCACCCGAACCTTCCACATCGAGGAGCCCCGCACAGGT
GGCCTGGCCAAGCGGCTGGCCAACCTTATCCGGCGCGCCCGAGGTGTCCCCCACTCT**TGA**TT
CTTGCCCTTTCCAGCAGCTGCAGCTGCCGTTTCTCTCTGGGGAGGGGAGCCCAAGGGCTGTT
TCTGCCACTTGCTCTCCTCAGAGTTGGCTTTTGAACCAAAGTGCCCTGGACCAGGTCAGGGC
CTACAGCTGTGTTGTCCAGTACAGGAGCCACGAGCCAAATGTGGCATTGAAATTTGAATTAA
CTTAGAAATTCATTTCTCACCTGTAGTGGCCACCTCTATATTGAGGTGCTCAATAAGCAAA
AGTGGTCGGTGGCTGCTGTATTGGACAGCACAGAAAAAGATTTCCATCACACAGAAAGGTC
GGCTGGCAGCACTGGCCAAGGTGATGGGGTGTGCTACACAGTGTATGTCAGTGTGTAGTGA
TGGAGTTTACTGTTTGTGGAATAAAAACGGCTGTTTCCGTGGAAAAA

238/330

FIGURE 238

MPLALLVLLLLGPGGWCLAEP PRDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHY
RLFPKALGQLISKYSLRELHLSFTQGFWRTRYWGPPFLQAPSGAELWVWFQDTVTDVDSWK
ELSNVLSGIFCASLNFIDSTNTVTPTASF KPLGLANDTDHYFLRYAVLPREVVCTENLTPWK
KLLPCSSKAGLSVLLKADRLEHTSYHSQAVHIRPVCARNARCTSI SWELRQTL SVVFDAFITG
QGKKDWSLFRMF SRTLTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTYQDVILGTRKT
YAIYDLLDTAMINNSRNLNIQLKWKRP PENEAPPVPFLHAQRYVSGYGLQKGELSTLLYNTH
PYRAFPVLLLDTPVWYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANS
VTKVSIQFERALLKWTEYTPDPNHGFYVSPSVLSALVPSMVAAPVDWEESPLFNSLFPVSD
GSNYFVRLYTEPLL VNLPTPDFSMPYNVICLTCTVAVCYGSFYNLLTRTFHIEEPRTGGLA
KRLANLIRRARGVPPL

239/330

FIGURE 239

CAAC**ATG**GGGTCCAGCAGCTTCTTGGTCCTCATGGTGTCTCTCGTTCTTGTGACCCTGGTGG
CTGTGGAAGGAGTTAAAGA⁵GGTATAGAGAAAGCAGGGGTTTGCCCAGCTGACAACGTACGC
TGCTTCAAGTCCGATCCTCCCCAGTGTCACACAGACCAGGACTGTCTGGGGGAAAGGAAGTG
TTGTTACCTGCACTGTGGCTTCAAGTGTGTGATTCTGTGAAGGAACTGGAAGAAGGAGGAA
ACAAGGATGAAGATGTGTCAAGGCCATACCCTGAGCCAGGATGGGAGGCCAAGTGTCCAGGC
TCCTCCTCTACCAGGTGTCCTCAGAAAT**TGA**TGCTGGGTCCTTTCTACCTCTGGGGGTCACTC
TCACTTGGCACCTGCCCCTGAGGGTCCTGAGACTTGGAATATGGAAGAAGCAATACCCAACC
CCACCAAAGAAAACCTGAGCTTGAAGTCCTTTTCCCCAAAAGAGGGAAGAGTCACAAAAAG
TCCAGACCCCAGGGACGGTACTTTCCCTCTCTACCTGGTGCTCCTCCCTAATGCTCATGAAT
GGACCCCTCATGAATGAAACCAGTGCCCTTATAAGAGACCCCAAAGAGCTGCCTTGCCCTTC
TGCAATGTGTGATCACAGCTAGAAGGCACTGTCAGAGAAGAGAAACTGGTCCTCACCAGATG
CTGAATCTGCTGGTGCCTTGATCTTGGACTTCCCAGCCTCTAGAACTGTAAGAAATAAATAT
TTGCTGTTTATAATCCAA

240/330

FIGURE 240

MGSSSFVLVLMVSLVLVTLVAVEGVKEGIEKAGVCPADNVRCFKSDPPQCHTDQDCLGERKCC
YLHCGFKCVIPVKELEEKGKDEDVSRPYPEPGWEAKCPGSSSTRCPQK

Signal sequence:

amino acids 1-19

N-myristoylation sites:

amino acids 23-29, 27-33, 32-38, 102-108

WAP-type 'four-disulfide core' domain signature:

amino acids 49-63

241/330

FIGURE 241

AAACTCAGCACTTGCCGGAGTGGCTCATTGTTAAGACAAAGGGTGTGCACTTCCTGGCCAGG
AAACCTGAGCGGTGAGACTCCCAGCTGCCTACATCAAGGCCCCAGGACATGCAGAACCTTCC
TCTAGAACCCGACCCACCACC**ATG**AGGTCTGCCTGTGGAGATGCAGGCACCTGAGCCAAGG
CGTCCAGTGGTCCTTGCTTCTGGCTGTCCTGGTCTTCTTTCTCTTCGCCTTGCCCTCTTTTA
TTAAGGAGCCTCAAACAAAGCCTTCCAGGCATCAACGCACAGAGAACATTAAAGAAAGGTCT
CTACAGTCCCTGGCAAAGCCTAAGTCCCAGGCACCCACAAGGGCGAGGAGGACAACCATCTA
TGCAGAGCCAGCGCCAGAGAACAAATGCCCTCAACACACAAACCCAGCCCAAGGCCACACCA
CCGGAGACAGAGGAAAGGAGGCCAACCAGGCACCGCCGGAGGAGCAGGACAAGGTGCCCCAC
ACAGCACAGAGGGCAGCATGGAAGAGCCCAGAAAAAGAGAAAACCATGGTGAACACACTGTC
ACCCAGAGGGCAAGATGCAGGGATGGCCTCTGGCAGGACAGAGGCACAATCATGGAAGAGCC
AGGACACAAAGACGACCCAAGGAAATGGGGGCCAGACCAGGAAGCTGACGGCCTCCAGGACG
GTGTCAGAGAAGCACCAGGGCAAAGCGGCAACCACAGCCAAGACGCTCATTCCCAAAAGTCA
GCACAGAATGCTGGCTCCACAGGAGCAGTGTCAACAAGGACGAGACAGAAAGGAGTGACCA
CAGCAGTCATCCCACCTAAGGAGAAGAAACCTCAGGCCACCCACCCCTGCCCTTTCCAG
AGCCCCACGACGCAGAGAAACCAAGACTGAAGGCCGCCAACTTCAAATCTGAGCCTCGGTG
GGATTTTGGAGAAAAATACAGCTTCGAAATAGGAGGCCTTCAGACGACTTGCCCTGACTCTG
TGAAGATCAAAGCCTCCAAGTCGCTGTGGCTCCAGAACTCTTTCTGCCCAACCTCACTCTC
TTCCTGGACTCCAGACACTTCAACCAGAGTGAGTGGGACCGCCTGGAACACTTTGCACCACC
CTTTGGCTTCATGGAGCTCAACTACTCCTTGGTGCAGAAGGTCGTGACACGCTTCCCTCCAG
TGCCCCAGCAGCAGCTGCTCCTGGCCAGCCTCCCCGCTGGGAGCCTCCGGTGCATCACCTGT
GCCGTGGTGGGCAACGGGGGCATCCTGAACAACCTCCCACATGGGCCAGGAGATAGACAGTCA
CGACTACGTGTTCCGATTGAGCGGAGCTCTCATTAAGGCTACGAACAGGATGTGGGGACTC
GGACATCCTTCTACGGCTTTACCGCCTTCTCCCTGACCCAGTCACTCCTTATATTGGGCAAT
CGGGGTTTCAAGAACGTGCCTCTTGGGAAGGACGTCCGCTACTTGCACTTCCTGGAAGGCAC
CCGGGACTATGAGTGGCTGGAAGCACTGCTTATGAATCAGACGGTGATGTCAAAAAACCTTT
TCTGGTTCAGGCACAGACCCAGGAAGCTTTTCGGGAAGCCCTGCACATGGACAGGTACCTG
TTGCTGCACCCAGACTTTCTCCGATACATGAAGAACAGGTTTCTGAGGTCTAAGACCCTGGA
TGGTGCCCACTGGAGGATATACCGCCCCACCACTGGGGCCCTCCTGCTGCTCACTGCCCTTC
AGCTCTGTGACCAGGTGAGTGCTTATGGCTTCATCACTGAGGGCCATGAGCGCTTTTCTGAT
CACTACTATGATACATCATGGAAGCGGCTGATCTTTTACATAAACCATGACTTCAAGCTGGA
GAGAGAAGTCTGGAAGCGGCTACACGATGAAGGGATAATCCGGCTGTACCAGCGTCCTGGTC
CCGGAACCTGCCAAAGCCAAGAAC**TGA**CCGGGGGCCAGGGCTGCCATGGTCTCCTTGCCTGCTC
CAAGGCACAGGATACAGTGGGAATCTTGAGACTCTTTGGCCATTTCCCATGGCTCAGACTAA
GCTCCAAGCCCTTCAGGAGTTCCAAGGGAACACTTGAACCATGGACAAGACTCTCTCAAGAT
GGCAAATGGCTAATTGAGGTTCTGAAGTTCTTCAGTACATTGCTGTAGGTCTGAGGCCAGG
GATTTTTAATTAAATGGGGTGATGGGTGGCCAATACCACAATTCCTGCTGAAAAACACTCTT
CCAGTCCAAAAGCTTCTTGATACAGAAAAAAGAGCCTGGATTTACAGAAACATATAGATCTG
GTTTGAATTCCAGATCGAGTTTACAGTTGTGAAATCTTGAAGGTATTACTTAACCTTCACTAC
AGATTGTCTAGAAGACCTTTCTAGGAGTTATCTGATTCTAGAAGGGTCTATACTTGTCTTG
TCTTTAAGCTATTTGACAACCTCTACGTGTTGTAGAAAACCTGATAATAATACAAATGATTGTT
GTCCATGGAAAGGCAAATAAATTTTCTACAGTGAAAAA

242/330

FIGURE 242

MRSCLRRCRHL SQGVQWSLL LAVLVFFLFALPSFIKEPQTKPSRHQRTENIKERSLQSLAKP
KSQAPTRARRTTIYAE PAPENNALNTQTQPKAHTTGDRGKEANQAPPEEQDKVPHTAQRAAW
KSPEKEKTMVNTLS PRGQDAGMASGRTEAQSWKSQDTKTTQGNNGGQTRKLTASRTVSEKHQG
KAATTAKTLIPKSQHRMLAPTGA VSTRTRQKGVTTAVIPPKEKKPQATPPPAPFQSPTTQRN
QRLKAANFKSEPRWDFEEKYSFEIGGLQTTCPDSVKIKASKSLWLQKLFLEPNLTLFLDSRHF
NQSEWDRLEHFAPPF GFMELNYSLVQKVVTFRFPVPQQQLLLASLPAGSLRCITCAVVGNGG
ILNNSHMGQEIDSHDYVFRLSGALIKGYEQDVGTRTSFYGFTAFSLTQSLILGNRGFKNVP
LGKDVRYLHFLEGTRDYEWLEALLMNQTVMSKNLFWFRHRPQEAFREALHMDRYLLLHPDFL
RYMKNRFLRSKTL DGAHWRIYRPTTGALLLLTALQLCDQVSAYGFITEGHERFSDHYDTSW
KRLIFYINHDFKLEREVWKRLHDEGIIRLYQRPGPGTAKAKN

Cytoplasmic Domain:

amino acids 1-10

Type II Transmembrane Domain:

amino acids 11-35

Lumenal catalytic Domain:

amino acids 36-600

Ribonucleotide Reductase small subunit Signature:

amino acids 481-496

N-glycosylation Sites:

amino acids 300-303, 311-314, 331-334, 375-378, 460-463

243/330

FIGURE 243

CG**ATG**CGCGGACCCGGGCACCCCCTCCTCCTGGGGCTGCTGCTGGTGCTGGGGCCTTCGCCG
GAGCAGCGAGTGGAAATTGTTCCCTCGAGATCTGAGGATGAAGGACAAGTTTCTAAAACACCT
TACAGGCCCTCTTTATTTTAGTCCAAAGTGCAGCAAACACTTCCATAGACTTTATCACAACA
CCAGAGACTGCACCATTCCCTGCATACTATAAAAGATGCGCCAGGCTTCTTACCCGGCTGGCT
GTCAGTCCAGTGTGCATGGAGGATAAG**TGA**GCAGACCGTACAGGAGCAGCACACCAGGAGCC
ATGAGAAGTGCCTTGGAACCAACAGGGAAACAGAACTATCTTTATACACATCCCCTCATGG
ACAAGAGATTTATTTTGCAGACAGACTCTTCCATAAGTCCTTTGAGTTTTGTATGTTGTTG
ACAGTTTGCAGATATATATTCGATAAATCAGTGTACTTGACAGTGTTATCTGTCACTTATTT

244/330

FIGURE 244

MRGPGHPLLLGLLLVLGPSPEQRVEIVPRDLRMKDKFLKHLTGPLYFSPKCSKHFHRLYHNT
RDCTIPAYYKRCARLLTRLAVSPVCMEDK

245/330

FIGURE 245

GGGCTGGGCCCCGCCGAGCTCCAGCTGGCCGGCTTGGTCCTGCGGTCCCTTCTCTGGGAGG
CCCGACCCCGGCCGCGCCCAGCCCCACC**ATG**CCACCCGCGGGGCTCCGCCGGGCCGCGCCG
CTCACCGCAATCGCTCTGTTGGTGCTGGGGGCTCCCCTGGTGCTGGCCGGCGAGGACTGCCT
GTGGTACCTGGACCGGAATGGCTCCTGGCATCCGGGGTTTAACTGCGAGTTCTTCACCTTCT
GCTGCGGGACCTGCTACCATCGGTACTGCTGCAGGGACCTGACCTTGCTTATCACCGAGAGG
CAGCAGAAGCACTGCCTGGCCTTCAGCCCCAAGACCATAGCAGGCATCGCCTCAGCTGTGAT
CCTCTTTGTTGCTGTGGTTGCCACCACCATCTGCTGCTTCCTCTGTTCTGTTGCTACCTGT
ACCGCCGGCGCCAGCAGCTCCAGAGCCCATTGAAGGCCAGGAGATTCCAATGACAGGCATC
CCAGTGCCAGCCAGTATACCCATACCCCAGGACCCCCAAAGCTGGCCCTGCACCCCCACAGCC
TGGCTTCATGTACCCACCTAGTGGTCCTGCTCCCCAATATCCACTCTACCCAGCTGGGCCCC
CAGTCTACAACCCTGCAGCTCCTCCTCCCTATATGCCACCACAGCCCTCTTACCCGGGAGCC
TGAGGAACCAGCCATGTCTCTGCTGCCCCCTTCAGTGATGCCAACCTTGGGAGATGCCCTCAT
CCTGTACCTGCATCTGGTCCTGGGGGTGGCAGGAGTCCTCCAGCCACCAGGCCCCAGACCAA
GCCAAGCCCTGGGCCCCTACTGGGGACAGAGCCCCAGGGAAGTGGAACAGGAGCTGAACTAGA
ACTATGAGGGGTTGGGGGGAGGGCTTGGAATTATGGGCTATTTTTACTGGGGGCAAGGGAGG
GAGATGACAGCCTGGGTCACAGTGCCTGTTTTCAAATAGTCCCTCTGCTCCCAAGATCCCAG
CCAGGAAGGCTGGGGCCCTACTGTTTGTCCCCTCTGGGCTGGGGTGGGGGGAGGGAGGAGGT
TCCGTCAGCAGCTGGCAGTAGCCCTCCTCTCTGGCTGCCCCACTGGCCACATCTCTGGCCTG
CTAGATTAAAGCTGTAAAGACAAAA

246/330

FIGURE 246

MPPAGLRRRAAPLTAIALLLVLGAPLVLAGEDCLWYLDNRNGSWHPGFNCEFFTFCCGTCYHRYC
CRDLTLLITERQQKHCLAFSPKTIAGIASAVILEFVAVVATTICCFLCSCCYLYRRRQQLQSP
FEGQEIPMTGIPVQPVYPYPQDPKAGPAPPQPGFMYPPSGPAPQYPLYPAGPPVYNPAAPPP
YMPPQPSYPGA

Transmembrane Domains:

amino acids 10-28, 85-110

N-glycosylation Site:

amino acids 38-41

N-myristoylation Sites:

amino acids 5-10, 88-93

247/330

FIGURE 247

GGGGGAGCTAGGCCGGCGGCAGTGGTGGTGGCGGCGGCGCAAGGGTGAGGGCGGCCCCAGAA
CCCCAGGTAGGTAGAGCAAGAAG**ATGGT**GTTTTCTGCCCTCAAATGGTCCCTTGCAACCATG
TCATTTCTACTTTCCCTCACTGTTGGCTCTCTTAACTGTGTCCACTCCTTCATGGTGTGAGAG
CACTGAAGCATCTCCAAAACGTAGTGATGGGACACCATTTCCTTGGAATAAAATACGACTTC
CTGAGTACGTATCCCAGTTCATTATGATCTCTTGATCCATGCAAACCTTACCACGCTGACC
TTCTGGGGAACCACGAAAGTAGAAATCACAGCCAGTCAGCCCACCAGCACCATCATCTGCA
TAGTCACCACCTGCAGATATCTAGGGCCACCCTCAGGAAGGGAGCTGGAGAGAGGCTATCGG
AAGAACCCCTGCAGGTCTGGAACACCCCCCTCAGGAGCAAATTGCACTGCTGGCTCCCGAG
CCCCCTCTGTGCGGGCTCCCGTACACAGTTGTCACTTCACTATGCTGGCAATCTTTCGGAGAC
TTTCCACGGATTTTACAAAAGCACCTACAGAACCAAGGAAGGGGAAGTGAAGGATACTAGCAT
CAACACAATTTGAACCCACTGCAGCTAGAATGGCCTTTCCCTGCTTTGATGAACCTGCCTTC
AAAGCAAGTTTCTCAATCAAATTAGAAGAGAGCCAAGGCACCTAGCCATCTCCAATATGCC
ATTGGTGAAATCTGTGACTGTTGCTGAAGGACTCATAGAAGACCATTTTGATGTCACTGTGA
AGATGAGCACCTATCTGGTGGCCTTCATCACTTTGAGATTTTGAGTCTGTCAGCAAGATAACC
AAGAGTGGAGTCAAGGTTTCTGTTTATGCTGTGCCAGACAAGATAAATCAAGCAGATTATGC
ACTGGATGCTGCGGTGACTCTTCTAGAATTTTATGAGGATTATTTGAGCATAACCGTATCCCC
TACCCAAACAAGATCTTGCTGCTATTCCCGACTTTCAGTCTGGTGTCTATGGAAAACCTGGGGA
CTGACAACATATAGAGAATCTGCTCTGTTGTTTGTATGTCAGAAAAGTCTTCTGCATCAAGTAA
GCTTGGCATCACAGTGACTGTGGCCCATGAACCTGGCCACCAGTGGTTTGGGAACCTGGTCA
CTATGGAATGGTGGAAATGATCTTTGGCTAAATGAAGGATTTGCCAAATTTATGGAGTTTGTG
TCTGTCACTGTGACCCATCCTGAACTGAAAGTTGGAGATTATTTCTTTGGCAAATGTTTTGA
CGCAATGGAGGTAGATGCTTTAAATTCCTCACACCCTGTGTCTACACCTGTGGAAAATCCTG
CTCAGATCCGGGAGATGTTTGATGATGTTTCTTATGATAAGGGAGCTTGTATTCTGAATATG
CTAAGGGAGTATCTTAGCGCTGACGCATTTAAAGTGGTATTGTACAGTATCTCCAGAAGCA
TAGCTATAAAAAATACAAAAACGAGGACCTGTGGGATAGTATGGCAAGTATTTGCCCTACAG
ATGGTGTAAAAGGGATGGATGGCTTTTGTCTAGAAAGTCAACATTCATCTTCATCCTCACAT
TGGCATCAGGAAGGGGTGGATGTGAAAACCATGATGAACACTTGGACACTGCAGAGGGGGTTT
TCCCCTAATAACCATCACAGTGAGGGGGAGGAATGTACACATGAAGCAAGAGCACTACATGA
AGGGCTCTGACGGCGCCCCGGACACTGGGTACCTGTGGCATGTTCCATTGACATTGATCACC
AGCAAATCCAACATGGTCCATCGATTTTTTGCTAAAAACAAAAACAGATGTGCTCATCCTCCC
AGAAGAGGTGGAATGGATCAAATTTAATGTGGGCATGAATGGCTATTACATTGTGCATTACG
AGGATGATGGATGGGACTCTTTGACTGGCCTTTTAAAAGGAACACACACAGCAGTCAGCAGT
AATGATCGGGCAAGTCTCATTAAACAATGCATTTGAGCTCGTCAGCATTTGGGAAGCTGTCCAT
TGAAAAGGCCTTGGATTTATCCCTGTACTTGAAACATGAAACTGAAATTATGCCCGTGTTTC
AAGGTTTGAATGAGCTGATTCCTATGTATAAGTTAATGGAGAAAAGAGATATGAATGAAGTG
GAAACTCAATTCAAGGCCTTCCTCATCAGGCTGCTAAGGGACCTCATTGATAAGCAGACATG
GACAGACGAGGGCTCAGTCTCAGAGCAAATGCTGCGGAGTGAACACTACTCCTCGCCTGTG
TGCACAACTATCAGCCGTGCGTACAGAGGGCAGAAGGCTATTTGAGAAAAGTGGAAAGGAATCC
AATGGAAAAGTTGAGCCTGCCTGTGACGTGACCTTGGCAGTGTTTGCTGTGGGGGCCAGAG
CACAGAAGGCTGGGATTTTCTTTATAGTAAATATCAGTTTTCTTTGTCCAGTACTGAGAAAA
GCCAAATTGAATTTGCCCTCTGCAGAACCCAAAATAAGGAAAAGCTTCAATGGCTACTAGAT
GAAAGCTTTAAGGGAGATAAAAATAAAAACCTCAGGAGTTTCCACAAATTCTTACACTCATTGG
CAGGAACCCAGTAGGATACCCACTGGCCTGGCAATTTCTGAGGAAAAACCTGGAACAAACTTG
TACAAAAGTTTGAACCTGGCTCATCTTCCATAGCCCACATGGTAATGGGTACAACAAATCAA
TTCTCCACAAGAACACGGCTTGAAGAGGTAAAAGGATTCTTCAGCTCTTTGAAAGAAAATGG
TTCTCAGCTCCGTTGTGTCCAACAGACAATTGAAACCATTTGAAGAAAACATCGGTTGGATGG
ATAAGAATTTTGATAAAATCAGAGTGTGGCTGCAAAGTGAAAAGCTTGAACGTATG**TAAAAA**
TTCTCCCTTGCCCGGTTCTCTGTTATCTCTAATACCAACATTTTGTGAGTGTATTTTCAA
ACTAGAGATGGCTGTTTTGGCTCCAACCTGGAGACTATTTTCCCTTCACTCATTTTGTGA
CTATCCCTGTGAAAAGAATAGCTGTTAGTTTTTCATGAATGGGCTTTTTTCATGAATGGGCTA
TCGCTACCATGTGTTTTGTTTCATCACAGGTGTTGCCCTGCAACGTAAACCCAAGTGTTGGGT
TCCCTGCCACAGAAGAATAAAGTACCTTATTCTCTCAAAAAAAAAAAAAAAAAAAAAAAAAA

248/330

FIGURE 248

MVFLPLKWSLATMSFLLSLLALLTVSTPSWCQSTEASPKRSDGTPFPWNKIRLPEYVIPVH
YDLLIHANLTTLTFWGTTKVEITASQPTSTIILHSHHLQISRATLRKGAGERLSEEPLQVLE
HPPQEQIALLAPEPLLVLGPYTVVIHYAGNLSETFHGFYKSTYRTKEGELRILASTQFEPTA
ARMAFPCFDEPAFKASFSIKIRREPRHLAISNMPLVKSVTVAEGLIEDHFDVTVKMSTYLVA
FIISDFESVSKITKSGVKVSVYAVPDKINQADYALDAAVTLLFYEYDYFSIPYPLPKQDLAA
IPDFQSGAMENWGLTTYRESALLFDAEKSSASSKLGITVTVAHELAHQWFGNLVTMEWWNDL
WLNEGFAKFMFVSVSVTHPELKVG DYFFGKCFDAMEVDALNSSHPVSTPVENPAQIREMFD
DVSYDKGACILNMLREYLSADAFKSGIVQYLQKHSYKNTKNEDLWDSMASICPTDGVKGM DG
FCRSRQSHSSSSSHWHQEGVDVKTMMNTWTLQRGFPLITITVRGRNVHMKQEHYMKGSDGAPD
TGYLWHVPLTFITSKSNMVHRFLLKTKTDVLILPEEVEWIKFNVGMNGYYIVHYEDDGWDSL
TGLLKGTHTAVSSNDRASLINNAFQLV SIGKLSIEKALDLSLYLKHETEIMPVFQGLNELIP
MYKLMEKRDMNEVETQFKAFILRLLRDLIDKQTTWTDEGSVSEQMLRSELLLLACVHNYQPCV
QRAEGYFRKWKESNGNLSLPVDVTLAVFAVGAQSTEGWDFLYSKYQFSLSSTEKSQIEFALC
RTQNKEKLQWLLDESFKGDKIKTQEFQILTIGRNPVGYPLAWQFLRKNWNKLVQKFELGS
SSIAHVMGTTNQFSTRTRLEE VKGFFSSLKENGSQLRCVQQTETIETIENIGWMDKNFDKIR
VWLQSEKLERM

Signal peptide:

amino acids 1-34

N-glycosylation sites:

amino acids 70-74, 154-158, 414-418, 760-764, 901-905

Neutral zinc metallopeptidases, zinc-binding region signature:

amino acids 350-360

249/330

FIGURE 249

CAGCCACAGACGGGTC**ATG**AGCGCGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCCAC
TGCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCC
GACCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGA
CACGTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGG
AGGCCAAGGACCAGGAGCCCCGCGTCACTGAGCACCGGATGGGCCCCGGCCTCTCCCTGATC
TCCTACACCTTCGTGTGCCGCCAGGAGGACTTCTGCAACAACCTCGTTAACTCCCTCCCGCT
TTGGGCCCCACAGCCCCCAGCAGACCCAGGATCCTTGAGGTGCCAGTCTGCTTGTCTATGG
AAGGCTGTCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGAT
GGCCTCCTCAGGCTCAGGGGAGGAGGCATCTTCTCCAATCTGAGAGTCCAGGGATGCATGCC
CCAGCCAGGTTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCCGTGGGTATGACTGAGA
ACTGCAATAGGAAAGATTTTCTGACCTGTATCGGGGGACCACCATTATGACACACGGAAAC
TTGGCTCAAGAACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGT
GTGTCAGGAGACGCTGCTGCTCATAGATGTAGGACTCACATCAACCTGGTGGGGACAAAAG
GCTGCAGCACTGTTGGGGCTCAAAATTTCCAGAAGACCACCATCCACTCAGCCCCCTCCTGGG
GTGCTTGTGGCCTCCTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAG
CAGCGTTCTGCTGAACTCCCTCCCTCCTCAAGCTGCCCCCTGTCCCAGGAGACCGGCAGTGTC
CTACCTGTGTGCAGCCCCCTTGGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGG
GGCGCCACTCATTGTTATGATGGGTACATTCATCTCTCAGGAGGTGGGCTGTCCACCAAAT
GAGCATTACAGGGCTGCGTGGCCCAACCTTCCAGCTTCTTGTGAACCACACCAGACAAATCG
GGATCTTCTCTGCGCGTGAGAAGCGTGATGTGCAGCCTCCTGCCTCTCAGCATGAGGGAGGT
GGGGCTGAGGGCCTGGAGTCTCTCACTTGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTG
GTGGGGAGTGGTTTGCCCTTCCTGC**TAA**CTCTATTACCCCCACGATTCTTCACCGCTGCTGA
CCACCCACACTCAACCTCCCTCTGACCTCATAACCTAATGGCCTTGAGACACCAGATTCTTTC
CCATTCTGTCCATGAATCATCTTCCCCACACACAATCATTCATATCTACTCACCTAACAGCA
ACACTGGGGAGAGCCTGGAGCATCCGGACTTGCCCTATGGGAGAGGGGACGCTGGAGGAGTG
GCTGCATGTATCTGATAATACAGACCCTGTCCTTTCA

250/330

FIGURE 250

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPRQWTPKNTSCDSGLGCQDTLMLI
ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP
PADPGSLRCPVCLSMEGCLEGTTEEICPKGTTHCYDGLLRRLRGGGIFSNLRVQGCMPQPGCN
LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSENTEMCEVGQVCQETL
LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN
SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGC
VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESLTWGVGLALAPALWWGVVC
PSC

251/330

FIGURE 251

GCGACGGGCAGGACGCCCCGTTCGCCTAGCGCGTGCTCAGGAGTTGGTGTCTGCCTGCGCT
CAGG**ATG**AGGGGGAATCTGGCCCTGGTGGGCGTTCTAATCAGCCTGGCCTTCCTGTCACTGCTG
CCATCTGGACATCCTCAGCCGGCTGGCGATGACGCCTGCTCTGTGCAGATCCTCGTCCCTGG
CCTCAAAGGGGATGCGGGAGAGAAGGGAGACAAAGGCGCCCCGGACGGCCTGGAAGAGTCG
GCCCCACGGGAGAAAAAGGAGACATGGGGGACAAAGGACAGAAAGGCAGTGTGGGTCGTCAT
GGAAAAATTGGTCCCATTTGGCTCTAAAGGTGAGAAAGGAGATTCCGGTGACATAGGACCCCC
TGGTCCTAATGGAGAACCAGGCCTCCCATGTGAGTGCAGCCAGCTGCGCAAGGCCATCGGGG
AGATGGACAACCAGGTCTCTCAGCTGACCAGCGAGCTCAAGTTCATCAAGAATGCTGTGCGC
GGTGTGCGCGAGACGGAGAGCAAGATCTACCTGCTGGTGAAGGAGGAGAAGCGCTACGCGGA
CGCCCAGCTGTCTGCCAGGGCCGCGGGGGCACGCTGAGCATGCCCAAGGACGAGGCTGCCA
ATGGCCTGATGGCCGCATACCTGGCGCAAGCCGGCCTGGCCCCGTGTCTTCATCGGCATCAAC
GACCTGGAGAAGGAGGGCGCCTTCGTGTACTCTGACCACTCCCCCATGCGGACCTTCAACAA
GTGGCGCAGCGGTGAGCCCAACAATGCCTACGACGAGGAGGACTGCGTGGAGATGGTGGCCT
CGGGCGGCTGGAACGACGTGGCCTGCCACACCACCATGTACTTCATGTGTGAGTTTGACAAG
GAGAACATG**TGA**GCCTCAGGCTGGGGCTGCCCATTTGGGGGCCCCACATGTCCCTGCAGGGTT
GGCAGGGACAGAGCCCAGACCATGGTGCCAGCCAGGGAGCTGTCCCTCTGTGAAGGGTGGAG
GCTCACTGAGTAGAGGGCTGTTGTCTAACTGAGAAAATGGCCTATGCTTAAGAGGAAAATG
AAAGTGTTCCTGGGGTGCTGTCTCTGAAGAAGCAGAGTTTCATTACCTGTATTGTAGCCCCA
ATGTCATTATGTAATTATTACCCAGAATTGCTCTTCCATAAAGCTTGTGCCTTTGTCCAAGC
TATACAATAAAATCTTTAAGTAGTGCAGTAGTTAAGTCCAAAAAAAAAAAAAAAAAAAAA

252/330

FIGURE 252

MRGNLALVGVLI SLAFLSLLPSGHPQPAGDDACSVQILVPGLKGDAGEKGDKGAPGRPGRVG
PTGEKGDMDKGQKGSVGRHGKIGPIGSKGEKGDSDIGPPGPNGEPGLPCECSQLRKAIGE
MDNQVSQLTSELKF I KNAVAGVRETESKIYLLVKEEKRYADAQLSCQGRGGTLSMPKDEAAN
GLMAAYLAQAGLARVFIGINDLEKEGAFVYSDHSPMRTFNKWRSGEPNNAYDEEDCVEMVAS
GGWNDVACHTTMYFMCEFDKENM

253/330

FIGURE 253

AGTGACTGCAGCCTTCCTAGATCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGCAG
CACCAGTGTGTGAGGGGAGCAGGCAGCGGTCCTAGCCAGTTCCTTGATCCTGCCAGACCACC
CAGCCCCCGGCACAGAGCTGCTCCACAGGCACC**ATG**AGGATCATGCTGCTATTCACAGCCAT
CCTGGCCTTCAGCCTAGCTCAGAGCTTTGGGGCTGTCTGTAAGGAGCCACAGGAGGAGGTGG
TTCCTGGCGGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTC
AAAAGCCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAA
GGAATCAACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGA
GGAGCGTCCAGCCAGAGGGAAAGACAGGACCTTTCTTACCTTCAGTGAGGGTTCCCTCGGCCC
CTTCATCCCAATCAGCTTGGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACC
TTT**TAA**GACTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGT
ATCCCCCGAGAGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTC
TTTCCCTGTCCCAATCCCCAGGTGCGCACGCTCCTGTTACCCTTTCTCTTCCCTGTTCTTGT
AACATTCTTGTGCTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAAACTGCA
TAGTGAATATCCCCAACCCCAATGGGCATTGACTGTAGAATACCCTAGAGTTCCTGTAGTGT
CCTACATTAAAAATATAATGTCTCTCTCTATTCCCTCAACAATAAAGGATTTTTGCATATGAA
AAA

254/330

FIGURE 254

MRIMLLFTAILAFSLAQSFQAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGLLK
ALSQASTDPKESTSPKRDMDHDFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLGSTGK
SSLGTEEQRPL

Important features:

Signal peptide:

amino acids 1-18

Tyrosine kinase phosphorylation site.

amino acids 36-45

N-myristoylation site.

amino acids 33-39, 59-65

Amidation site.

amino acids 90-94

Leucine zipper pattern.

amino acids 43-65

Tachykinin family signature.

amino acids 86-92

255/330

FIGURE 255

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGCTGTGCCCCGCTGTGCCTGCTGTGCC
CGCGCTGTGCGCGCTGCTACCGCGTCTGCTGGACGCGGGAGACGCCAGCGAGCTGGTGATTG
GAGCCCTGCGGAGAGCTCAAGCGCCCAGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTC
CCATAGTTGCTGCAGGAGTGGAGCC**ATG**AGCTGCGTCCTGGGTGGTGTTCATCCCCCTTGGGGC
TGCTGTTCCCTGGTCTGCGGATCCCAAGGCTACCTCCTGCCCAACGTCACTCTCTTAGAGGAG
CTGCTCAGCAAATACCAGCACAAACGAGTCTCACTCCCGGGTCCGCAGAGCCATCCCCAGGGA
GGACAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGCCAGGTGCAGCCTCAGGCCT
CCAACATGGAGTACATGGTGAGCGCCGGCTCCGGCCGCAGAGGCTGGCACCGGGGGTGGGGC
CTGGGGCCACCAGCCTGCTCTGTTCCCCAGCCAGCTCTGTTCCCCAGCCAGTGCGTGTGATGG
CTGGCTCAGGGTCTCCTCTGGCAGGGGAGGATCCCGGCTCTGTTCTGTTTTGTTTGTGTTGTT
TTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAATCGTCATGCCCTG
AAACCT**TAG**ACTCCCGGGGTTAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAACTACAG
GCATGCACCATGGTGCCAGCTAGATTTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGT
TGCCCAGGCTGGTCTTGAACTCCTAGGCTCAAGCAATCCTCCTGCCTCAGCCTCTCAAAGTG
CTAGGATTATAGGCATGAGTCACCCTGTCTGGCTCTGGCTCTGTTCTTAACATTCTGCCAAA
ACAACACACGTGGGTTCCTGTGCAGAGCCTGCCTCGTTGCCTTCATGTCACTCTTGGTAGC
TCCACTGGGAACACAGCTCTCAGCCTTTCCCACCTGGAGGCAGAGTGGGGAGGGGGCCCAGGG
CTGGGCTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTC
CTTAGCCCGTGTGAGCCTCACTTTCCACTTGGAGAGTCCTTCCTCGCGTGGTTGCCATGACT
GTGAGATAAGTCGAGGCTGTGAAGGGCCCGGCACAGACTGACCTGCCTCCCCAACCCCTAGG
CTTTGCTAACCGGGAAAGGAGCTAACGGTGACAGAAGACAGCCAAGGTCAACCCTCCCGGGT
GATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGACCCCAAGCTCCAGTG
TGGAACCTTCCTTCCTGGCTGGTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAGG
GTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTCAGCTATGAATGGCTT
TTTAAACAAACCCACGTCCCAGCCTGGGTAACATGGTAAAGCCCCGTCTCTACAAAAAATC
CAAGTTAGCCGGGCATGGTGGTGCGCACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAG
GTGGAGGTGGGGGGTGGGAGCTGAGGAAGGAGGATCGCTTGAGCCTGGGAAGTCGAGGCTGC
AGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCTGTCTCAAAAA

256/330

FIGURE 256

MSCVLGGV IPLGLLFLVCGSQGYLLPNVTLLLEELLSKYQHNEHSRVRRAIPREDKEEILML
HNKLRGQVQPQASNMEYMVSAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGWLRVSSGR
GGSRLCSVLFVCFETGSHSATDAGVQWHNRHALKP

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 27-31, 41-45

N-myristoylation site.

amino acids 126-132, 140-146

Amidation site.

amino acids 85-89

257/330

FIGURE 257

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACT**ATGGG**
GTCTGGGCTGCCCCCTTGTCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGG
GTATGACTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGC
TTCCTGGAATTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCAC
CCTCCACCATGCAAGATCTCAACACCATGTTGTCTGCAACACA**TGA**CAGCCATTGAAGCCTG
TGTCTTCTTGGCCCGGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTT
CAGCAGGCCCCCACCTCCTGAGTGGCAATAAATAAAATTCGGTATGCTG

258/330

FIGURE 258

MGSGPLVLLLTLGSSHGTGPGMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSGTS
VTLHHARSQHHVVCNT

259/330

FIGURE 259

AATTGTATCTGTGTAATGTTAAAACAAACGAAATAAAATAGAAGGAAAACTTTCTGAGTTT
CAAAAACAACAGACTAGTACTCTAAAGAACTCTTTAAAACAATTAAGTGTAGGATTGCAGT
TATGATTGGATATTATTTAATTCTGTTTCTGATGTGGGGTTCCTCCACTGTGTTCTGTGTGC
TATTAATATTTACCATTGCAGAAGCTTCATTCAGTGTTGAAAATGAATGCTTAGTGGATCTG
TGCCTCTTACGCATATGTTACAAATTATCTGGAGTTCCTAATCAATGCAGAGTTCCCTCCC
CTCCGATTGTTCTAAAT**TAA**ATTGAAAGATGTCTGCTGTGGAAAAAGGCATGTATTTAAATCTG
TATGATTCTCAACCATCTTTAGTTGGGAAAGGTCCTTGAAAGCCAATGGAAATACTTTTTTT
TTTTCTTGGCACTAATCAAGTGAGTGTTACCTTTTCACTTAGTAGGATGTGTTGTTACGCTA
GTAAAATAGAAACCTGTGTTTATTCTCAGGTATTTTAGAAACAACAGCCATCATTTTATTTT
ATGTGTGTGTTCTTGGCTGTATTCATAAATTATATATTTTGGGCTATCAAATATTACTTCAT
TCAATATAAATAACAATAGTAGAAGTTGTTTACTTAGATATGCTTTCTAGTTGCATTTTCTC
AGCCTATGTAAGACTACTTTGTTGTAATAGCCTTTGAAATTTACAGTACTGTCTCTCTACTA
TCTTCAGATTACTTGATTCAAATAAACCAATTATGTTTGTAATTGATATTAATAAAACCAGA
ATAAAAGTTCATATCTACCC

260/330

FIGURE 260

MIGYYLILFLMWGSSTVFCVLLIFTIAEASFVENECLVDLCLLRICYKLSGVPNQCRVPLP
SDCSK

Important features:

Signal peptide:

amino acids 1-29

261/330

FIGURE 261

GAGGATTTGCCACAGCAGCGGATAGAGCAGGAGAGCACCACCGGAGCCCTTGAGACATCCTT
GAGAAGAGCCACAGCATAAGAGACTGCCCTGCTTGGTGTTTTGCAGG**ATG**ATGGTGGCCCTT
CGAGGAGCTTCTGCATTGCTGGTTCTGTTCCCTTGCAGCTTTTCTGCCCCCGCCGAGTGTA
CCAGGACCCAGCCATGGTGCATTACATCTACCAGCGCTTTCGAGTCTTGGAGCAAGGGCTGG
AAAAATGTACCCAAGCAACGAGGGGCATACATTCAAGAATTCOAAGAGTTCTCAAAAAATATA
TCTGTCTATGCTGGGAAGATGTCAGACCTACACAAGTGAGTACAAGAGTGCAGTGGGTAACTT
GGCACTGAGAGTTGAACGTGCCCAACGGGAGATTGACTACATACAATACCTTCGAGAGGCTG
ACGAGTGCATCGTATCAGAGGACAAGACACTGGCAGAAATGTTGCTCCAAGAAGCTGAAGAA
GAGAAAAAGATCCGGACTCTGCTGAATGCAAGCTGTGACAACATGCTGATGGGCATAAAGTC
TTTGAATAAGTGAAGAAGATGATGGACACACATGGCTCTTGGATGAAAGATGCTGTCTATA
ACTCTCCAAAGGTGTACTTATTAATTGGATCCAGAAACAACACTGTTTGGGAATTTGCAAAC
ATACGGGCATTCATGGAGGATAACACCAAGCCAGCTCCCCGGAAGCAAATCCTAACACTTTC
CTGGCAGGGAACAGGCCAAGTGATCTACAAAGGTTTTCTATTTTTTTCATAACCAAGCAACTT
CTAATGAGATAATCAAATATAACCTGCAGAAAGAGGACTGTGGAAGATCGAATGCTGCTCCCA
GGAGGGGTAGGCCGAGCATTGGTTTACCAGCACTCCCCCTCAACTTACATTGACCTGGCTGT
GGATGAGCATGGGCTCTGGGCCATCCACTCTGGGCCAGGCACCCATAGCCATTTGGTTCTCA
CAAAGATTGAGCCGGGCACACTGGGAGTGGAGCATTTCATGGGATACCCCATGCAGAAGCCAG
GATGCTGAAGCCTCATTCCCTCTTGTGTGGGGTTCTCTATGTGGTCTACAGTACTGGGGGCCA
GGGCCCTCATCGCATCACCTGCATCTATGATCCACTGGGCACTATCAGTGAGGAGGACTTGC
CCAACTTGTCTTCCCCAAGAGACCAAGAAGTCACTCCATGATCCATTACAACCCAGAGAT
AAGCAGCTCTATGCCTGGAATGAAGGAAACCAGATCATTACAAACTCCAGACAAAGAGAAA
GCTGCCTCTGAAG**TAA**TGCATTACAGCTGTGAGAAAGAGCACTGTGGCTTTGGCAGCTGTTC
TACAGGACAGTGAGGCTATAGCCCCTTCACAATATAGTATCCCTCTAATCACACACAGGAAG
AGTGTGTAGAAGTGGAATACGTATGCCTCCTTTCCCAAATGTCACTGCCTTAGGTATCTTC
CAAGAGCTTAGATGAGAGCATATCATCAGGAAAGTTTCAACAATGTCCATTACTCCCCCAA
CCTCCTGGCTCTCAAGGATGACCACATTCTGATACAGCCTACTTCAAGCCTTTTGTTTTACT
GCTCCCCAGCATTTACTGTAACCTGCCATCTTCCCTCCCACAATTAGAGTTGTATGCCAGC
CCCTAATATTACCACTGGCTTTTCTCTCCCCTGGCCTTTGCTGAAGCTCTTCCCTCTTTTT
CAAATGTCTATTGATATTCTCCCATTTTCACTGCCCAACTAAAATACTATTAATATTTCTTT
CTTTTCTTTTCTTTTTTTTTGAGACAAGGTCTCACTATGTTGCCAGGCTGGTCTCAAACCTCC
AGAGCTCAAGAGATCCTCCTGCCTCAGCCTCCTAAGTACCTGGGATTACAGGCATGTGCCAC
CACACCTGGCTTAAAATACTATTTCTTATTGAGGTTTAACTCTATTTCCCTAGCCCTGTC
CTTCCACTAAGCTTGGTAGATGTAATAATAAAGTGAAAAATATTAACATTTGAATATCGCTTT
CCAGGTGTGGAGTGTGTTGCACATCATTGAATTCTCGTTTCACCTTTGTGAAACATGCACAAG
TCTTTACAGCTGTCATTCTAGAGTTTAGGTGAGTAACACAATTACAAAGTGAAAGATACAGC
TAGAAAATACTACAAATCCCATAGTTTTTCCATTGCCCAAGGAAGCATCAAATACGTATGTT
TGTTACCTACTCTTATAGTCAATGCGTTCATCGTTTCAGCCTAAAAATAATAGTCTGTCCC
TTTAGCCAGTTTTTCATGTCTGCACAAGACCTTTCAATAGGCCTTTCAAATGATAATTCCTCC
AGAAAACCAGTCTAAGGGTGAGGACCCCAACTCTAGCCTCCTCTTGTCTTGCTGTCTCTGT
TTCTCTCTTTCTGCTTTAAATTCAATAAAAGTGACACTGAGCAAAAAAAAAAAAAA

262/330

FIGURE 262

MMVALRGASALLVLFLLAAFLPPPQCTQDPAMVHYIYQRFVLEQGLEKCTQATRAYIQEFQE
FSKNISVMLGRCQTYTSEYKSAVGNLALRVERAQREIDYIQYLREADECIVSEDKTLAEMLL
QEAEKKKIRTLNASCNDMLMGIKSLKIVKKMMDTHGSWMKDAVYNBPKVYLLIGSRNNTV
WEFANIRAFMEDNTKPAPRKQILTLWQGTGQVIYKGFLFFHNQATSNEIIKYNLQKRTVED
RMLLPGGVGRALVYQHSPSTYIDLAVDEHGLWAIHSGPGTHSHLVLTKEPGLGVEHSWDT
PCRSQDAEASFLLCGVLYVVYSTGGQGPHRITCIYDPLGTISEEDLPNLFFPKRPRSHSMIH
YNPRDKQLYAWNENQIIYKLQTKRKLPLK

FIGURE 263

[illegible]

264/330

FIGURE 264

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQK
QYQIRIRKEKPQQHNFTHRLLAAALKSHSGNISCMDFSSNGKYLATCADDRRTIRIWSTKDFLQ
REHRSMRANVELDHATLVRFSPPDCRAFIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPKK
HKAPVIDIGIANTGKFIMTASSDTTVLIWSLKGQVLSTINTNQMNNTAAVSPCGRFVASC
GFTPDVKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKDGTWKLWDTDV
EYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVH
GECIANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHKLRASNESTRQRLQQQLTQ
AQETLKSLGALKK

Important features:**Signal peptide:**

amino acids 1-25

N-glycosylation site.

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

Beta-transducin family Trp-Asp repeat protein.

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

265/330

FIGURE 265

TGGCCTCCCCAGCTTGCCAGGCACAAGGCTGAGCGGGAGGAAGCGAGAGGCATCTAAGCAGG
CAGTGTTTTGCCTTCACCCCAAGTGACC**ATG**AGAGGTGCCACGCGAGTCTCAATCATGCTCC
TCCTAGTAACTGTGTCTGACTGTGCTGTGATCACAGGGGCCTGTGAGCGGGATGTCCAGTGT
GGGGCAGGCACCTGCTGTGCCATCAGCCTGTGGCTTCGAGGGCTGCGGATGTGCACCCCGCT
GGGGCGGGAAGGCGAGGAGTGCCACCCCGGCAGCCACAAGGTCCCCTTCTTCAGGAAACGCA
AGCACCACACCTGTCCTTGCTTGCCCAACCTGCTGTGCTCCAGGTTCCCGGACGGCAGGTAC
CGCTGCTCCATGGACTTGAAGAACATCAATTTT**TAG**GCGCTTGCCCTGGTCTCAGGATACCCA
CCATCCTTTTCCTGAGCACAGCCTGGATTTTTATTCTGCCATGAAACCCAGCTCCCATGAC
TCTCCAGTCCCTACACTGACTACCTGATCTCTCTTGTCTAGTACGCACATATGCACACAG
GCAGACATACCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGG
CTGTGGTGTGAAAGGTGGCCAGCCTGGTTCTCTTCCCTGCTCAGGCTGCCAGAGAGGTGGTA
AATGGCAGAAAGGACATTCCCCCTCCCCTCCCCAGGTGACCTGCTCTCTTTCCTGGGCCCTG
CCCCTCTCCCCACATGTATCCCTCGGTCTGAATTAGACATTCTTGGGCACAGGCTCTTGGGT
GCATTGCTCAGAGTCCCAGGTCCTGGCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGG
ACCAATTTGTGGGTAGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGAC
TCAAGATTGGCTCTTCCCAGAGGGCAGCAGACAGTCACCCCAAGGCAGGTGTAGGGAGCCCA
GGGAGGCCAATCAGCCCCCTGAAGACTCTGGTCCCAGTCAGCCTGTGGCTTGTGGCCTGTGA
CCTGTGACCTTCTGCCAGAATTGTCATGCCTCTGAGGCCCTCTTACCACACTTTACCAGT
TAACCACTGAAGCCCCCAATTCCCACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAA
TCTAATCTGATATTGACATATTAGAAGGCAATTAGGGTGTTTCCTTAAACAACCTCCTTTCCA
AGGATCAGCCCTGAGAGCAGGTGGTGACTTTGAGGAGGGCAGTCCTCTGTCCAGATTGGGG
TGGGAGCAAGGGACAGGGAGCAGGGCAGGGGCTGAAAGGGGCACTGATTACAGACCAGGGAGG
CAACTACACACCAACATGCTGGCTTTAGAATAAAAGCACCAACTGAAAAA

266/330

FIGURE 266

MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAISLWLRGLRMCTPLGREGEECHP
GSHKVPFFFRKRKHHTCPCLPNLLCSRFPDGRYRCSMDLKNINF

Signal peptide:

amino acids 1-19

Tyrosine kinase phosphorylation site:

amino acids 88-95

N-myristoylation sites:

amino acids 33-39, 35-41, 46-52

267/330

FIGURE 267

AGCGCCCGGGCGTCGGGGCGGTAAAAGGCCGGCAGAAGGGAGGCACTTGAGAA**ATG**TCTTTC
CTCCAGGACCCAAGTTTCTTCACCATGGGGATGTGGTCCATTGGTGCAGGAGCCCTGGGGGC
TGCTGCCTTGGCATTGCTGCTTGCCAACACAGACGTGTTTCTGTCCAAGCCCCAGAAAGCGG
CCCTGGAGTACCTGGAGGATATAGACCTGAAAACACTGGAGAAGGAACCAAGGACTTTCAAA
GCAAAGGAGCTATGGGAAAAAAATGGAGCTGTGATTATGGCCGTGCGGAGGCCAGGCTGTTT
CCTCTGTGCGAGAGGAAGCTGCGGATCTGTCCTCCCTGAAAAGCATGTTGGACCAGCTGGGCG
TCCCCCTCTATGCAGTGGTAAAGGAGCACATCAGGACTGAAGTGAAGGATTTCCAGCCTTAT
TTCAAAGGAGAAATCTTCCTGGATGAAAAGAAAAAGTTCTATGGTCCACAAAGGCGGAAGAT
GATGTTTATGGGATTTATCCGTCTGGGAGTGTGGTACAACCTTCTTCCGAGCCTGGAACGGAG
GCTTCTCTGGAAACCTGGAAGGAGAAGGCTTCATCCTTGGGGGAGTTTTCGTGGTGGGATCA
GGAAAGCAGGGCATTTCTTCTTGAGCACCGAGAAAAAGAATTTGGAGACAAAGTAAACCTACT
TTCTGTTCTGGAAGCTGCTAAGATGATCAAACCACAGACTTTGGCCTCAGAGAAAAAA**TGAT**
TGTGTGAAACTGCCCAGCTCAGGGATAACCAGGGACATTCACCTGTGTTTCATGGGATGTATT
GTTTCCACTCGTGTCCCTAAGGAGTGAGAAACCCATTTATACTCTACTCTCAGTATGGATTA
TTAATGTATTTTAATATTCTGTTTAGGCCCACTAAGGCAAAATAGCCCCAAAACAAGACTGA
CAAAAATCTGAAAAACTAATGAGGATTATTAAGCTAAAACCTGGGAAATAGGAGGCTTAAAA
TTGACTGCCAGGCTGGGTGCAGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAAGG
TGAGCAAGTCACTTGAGGTCGGGAGTTCGAGACCAGCCTGAGCAACATGGCGAAACCCCGTC
TCTACTAAAAATACAAAATCACCCGGGTGTGGTGGCAGGCACCTGTAGTCCCAGCTACCCG
GGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGAGGTGGAGGTTGCGGTGAGCTGAGATCA
CACCACTGTATTCCAGCCTGGGTGACTGAGACTCTAACTAA

268/330

FIGURE 268

MSFLQDPSFFTGMWSIGAGALGAAALALLANTDVFLSKPQKALEYLEDIDLKTLEKEPR
TFKAKELWEKNGAVIMAVRRPGCFLCREEAADLSSLKSMLDQLGVPLYAVVKEHIRTEVKDF
QPYFKGEIFLDEKKKFYGPQRRKMMFMGFIRLGVWYNFFRAWNGGFSGNLEGEFGFILGGVFV
VGSGKQGILLEHREKEFGDKVNLLSVLEAAKMIKPQTLASEKK

269/330

FIGURE 269

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCG
GGCCAGGTGCCCCGTGCGAGGTGCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGA
AGCCCCTTCCTCGGCGCTGCCAACCCGCCACCCAGCCC**ATG**GCGAACCCCGGGCTGGGGCTG
CTTCTGGCGCTGGGCCTGCCGTTCTTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACA
GACCACTTCTGCAAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATG
GCAACCTGCGTCCGGAAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTG
CTCCTGGCTGTGGGGCTGGCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGG
CACCTACCGGCCCAGTAGCGAGGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGG
ACTCCAAGGAGACGGTGCAGGGCTGCCTGCCCATC**TAG**GTCCCCTCTCCTGCATCTGTCTCC
CTTCATTGCTGTGTGACCTTGGGGAAAGGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAG
TGCTTAATAGCAGGGAAGAAGGTACTTCAAAGACTCTGCCCCTGAGGTCAAGAGAGGATGGG
GCTATTCACCTTTTATATATTTATATAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

270/330

FIGURE 270

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSSSTSSSSDGNLRPEAITAIIV
VFSLLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

271/330

FIGURE 271

AATATATCATCTATTTATCATTAATCAATAATGTATTCTTTTATTCCAATAACATTTGGGTT
TTGGGATTTTAAATTTTCAAACACAGCAGAAATGACATTTTTTCTGTCACTATTATTATTGTTG
GTATGTGAAGCTATTTGGAGATCCAATTCAGGAAGCAACACATTGGAGAATGGCTACTTTCT
ATCAAGAAATAAAGAGAACCACAGTCAACCCACACAATCATCTTTAGAAGACAGTGTGACTC
CTACCAAAGCTGTCAAAACCACAGGCAAGGGCATAGTTAAAGGACGGAATCTTGACTCAAGA
GGGTTAATTCTTGGTGCTGAAGCCTGGGGCAGGGGTGTAAAGAAAACACTTAGATTCAATG
ATTGTAAATTTAAGGCAAATACACATATTAGTATTACCTTAGTGTAATGTATCCCTGTCATA
TATACAATAAGGTGAAATTATAAGTACCCTATGCAGTTGGCTGGACAGTTCTAAATTGGACT
TTATTAATTTTTTAAATCAGTAACTGATTTATCACTGGCTATGTGCTTAGATCTACAGGAGA
TCATATAATTTGATACAAATAAAAGAAAAGTGTCTCTCCCCTTACAGAATTGACATTTTAA
ATGCGATACAGTTAGAATAGGAAATATGACATTAGAAAGGAAGAATGACAGGGAGAAAGGAA
AGAAGGGAAAATGTTGCCAAGGAAAAAAAAA

272/330

FIGURE 272

MTFFLSLLLLLVCEAIWRSNSGSNTLENGYFLSRNKENHSQPTQSSLEDSVTPTKAVKTTGK
GIVKGRNLDSRGLILGAEAWGRGVKKNT

FIGURE 273

GCCAGGGAATAACTAGAGAGGGAACA**ATG**GGGTTATTTCAGAGGTTTTGTTTTCTCTTAGTTCT
GTGCCTGCTGCACCAGTCAAATACTTCTTCATTAAAGCTGAATAATAATGGCTTTGAAGATA
TTGTCATTGTTATAGATCCTAGTGTGCCAGAAGATGAAAAAATAATTGAACAAATAGAGGAT
ATTGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAAGATTTTTTTTCAAAAA
TGTATCTATATTAATTCTCTGAGAAATTGGAAGGAAAAATCCTCAGTACAAAAGGCCAAAAATG
AAAACCATAAACATGCTGATGTTATAGTTGCACCACCTACACTCCCAGGTAGAGATGAACCA
TACACCAAGCAGTTTCACAGAATGTGGAGAGAAAAGCGAATACATTCACTTCACCCCGTACCT
TCTACTTTGGAAAAAAACAAAAATGAATATGGACCACCAAGGCAAACCTGTTTGCTCATGAGTGGG
CTCACCTCCGGTGGGGAGTGTTGATGAGTACAATGAAGATCAGCCTTTCTACCGTGCTAAG
TCAAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAA
GTGTCAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAATTGATTCTACAACAAAACCTGTATG
GAAAAGATTGTCAATTCTTTCTGTATAAAGTACAAACAGAAAAAGCATCCATAATGTTTATG
CAAAGTATTGATTCTGTTGTAATTTTGTAAACGAAAAACCCATAATCAAGAAGCTCCAAG
CCTACAAAACATAAAGTGCAATTTTGAAGTACATGGGAGGTGATTAGCAATTTCTGAGGATT
TTAAAAACACCATAACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATC
AGTCAAAGAATTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCCT
AAATCGAATGAATCAAGCAGCAAAAACATTTCTGTCTGCAGACTGTTGAAAATGGATCCTGGG
TGGGGATGGTTCATTTGATAGTACTGCCACTATTGTAATAAGCTAATCCAAATAAAAAGC
AGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATCTCTCTGGGAGGAACCTCCAT
CTGCTCTGGAATTAATATGCATTTTCAGGTGATTGGAGAGCTACATTCCCAACTCGATGGAT
CCGAAGTACTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTTCTTGTATTGATGAAGTG
AAACAAAGTGGGGCCATTGTTTCATTTTATTTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAAT
AGAGATGAGCAAGATAACGAGGAGGAAGTCATTTTGTATTGTTTCAGATGAAGCTCAGAACAATG
GCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAATACTGATCTCTCCAGAAAGTCCCTT
CAGCTCGAAAGTAAGGGATTAACACTGAATAGTAATGCCTGGATGAACGACACTGTCATAAT
TGATAGTACAGTGGGAAAGGACACGTTCTTTCTCATCACATGGAACAGTCTGCCTCCCAGTA
TTTCTCTCTGGGATCCCAGTGGAAACAATAATGGAAAATTTACAGTGGATGCAACTTCCAAA
ATGGCTCTATCTCAGTATTCCAGGAAGTCGAAAGGTGGGCACCTGGGCATACAATCTTCAAGC
CAAAGCGAAACCCAGAAACATTAACATATTACAGTAACCTCTCGAGCAGCAAATTTCTTCTGTG
CTCCAATCACAGTGAATGCTAAAAATGAATAAGGACGTAAACAGTTTCCCCAGCCCCAATGATT
GTTTACGCAGAAATTTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTCAT
TGAATCACAGAATGGACATACAGAAGTTTTTGGAACTTTTGGGATAATGGTGCAAGGCTGATT
CTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATATACAGAAAAATGGCAGATAT
AGCTTAAAAAGTTTCGGGCTCATGGAGGAGCAAACACTGCCAGGCTAAAATTTACGGCCTCCACT
GAATAGAGCCGCGTACATACCAGGCTGGGTAGTGAACGGGGAAATTTGAAGCAAACCCGCCAA
GACCTGAAATTGATGAGGATACTCAGACCACCTTGGAGGATTTACAGCCGAACAGCATCCGGA
GGTGCAATTTGTGGTATCACAAAGTCCCAAGCCTTCCCTTGCTGACCAATACCCACCAAGTCA
AATCACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTTCTTACATGGACAGCACGAG
GAGATAATTTTGATGTTGGAAAAGTTCAACGTTATATCATAAGAATAAGTGCAAGTATTCTT
GATCTAAGAGACAGTTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTCACCAAAGGA
GGCCAACTCCAAGGAAAGCTTTGCATTTAAACCAGAAAAATATCTCAGAAGAAAAATGCAACCC
ACATATTTATTGCCATTTAAAGATATAGATAAAAGCAATTTGACATCAAAAGTATCCAAACATT
GCACAAGTAACCTTTGTTTATCTCTCAAGCAAATCCTGATGACATTGATCCTACACCTACTCC
TACTCCTACTCCTACTCCTGATAAAAGTCAATAATCTGGAGTTAATATTTCTACGCTGGTAT
TGTCTGTGATTGGGTCTGTTGTAATTGTTAACTTTATTTTAAAGTACCACCATT**TGA**ACCTTA
ACGAAGAAAAAAATCTTCAAGTAGACCTAGAAGAGAGTTTTTAAAAAACAAAACAAATGTAAGT
AAAGGATATTTCTGAATCTTAAATTCATCCCATTGTGATCATAAATCATATAAAATAATTT
TTAAGATGTCGGAAGAGGATCTTTGATTAAATAAAAAACACTCATGGATATGTAAAAACTGT
CAAGATTAAAAATTTAATAGTTTTCATTTATTTGTTATTTTATTTGTAAGAAATAGTGATGAAC
AAAGATCCTTTTTTCATACTGATACCTGGTTGTATATTATTTGATGCAACAGTTTTCTGAAAT
GATATTTCAAATTGATCAAGAAATTAATAATCATCTATCTGAGTAGTCAAAATACAAGTAA
GGAGAGCAAAATAAACAACTTTGAAAAAAAATAAAAAATAAAAAATAAAAAATAAAAA
AAAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAA

274/330

FIGURE 274

MGLFRGFVFLVLCLLHQSNSTSFIKLNNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTTASTY
LFEATEKRFFFKNVSILIPENWKENPQYKRPKHENHKKHADVIVAPPTLPGRDEPYTKQFTEC
GEKGEYIHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPPFYRAKSKKIEATR
CSAGISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDKVQTEKASIMFMQSIDSVVE
FCNEKTHNQEAPSLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLKISQRIVCLV
LDKSGSMGGKDRLNRMNQAAKHFLQLTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLM
AGLPTYPLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLTGDGENTASSCIDEVKQSGAIVH
FIALGRAADEAVIEMSKITGGSHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLT
LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPG
TAKVGTWAYNLQAKANPETLTITVTSRAANSSVPPITVNAKMNKDVNSFPSPMIVYAEILQG
YVPVLGANVTAFIESQNGHTEVLELLDNGAGADSFKNDGVYSRYFTAYTENGRYSLKVBRAHG
GANTARLKLRPPLNRAAYIPGWVNGEIEANPPRPEIDEDTQTTLEDFSRASGGAFVVSQV
PSLPLPDQYPPSQITDLDATVHEDKIILTWTAPGDNFDVGKVQRYIIRISASILDLRDSFDD
ALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAIKSIDKSNLTSKVSNIAQVTLFIP
QANPDDIDPTPTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIVNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

275/330

FIGURE 275

CTCCTTAGGTGGAAACCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAAGACCATACGTCCCCG
GGCAGGGGTGACAACAGGTGTCTCTTTTGTATCTCGTGTGTGGCTGCCTTTCCTATTTCAAGGAAAG
ACGCCAAGGTAAATTTTGACCCAGAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACC
CCCAGTTATGCCAGGATTTACTAGAGAGTGTCAACTCAACCAGCAAGCGGCTCCTTCGGCTTAACCT
GTGGTTGGAGGAGAGAACCCTTGTGGGGCTGCGTTCTCTTAGCAGTGCTCAGAAAGTGAAGTTCGCTGA
GGGTGGACCAGAAGAAAGGAAAGGTCCCCCTCTTGCTGTTGGCTGCACATCAGGAAGGCTGTGATGGG
AATGAAGGTGAAAACCTTGAGATTTTCACTTTCAGTTCATTGCTTCTGCCTGCAAGATCATCCTTTAAAA
GTAGAGAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATG
CAAGCAGCTCCGGGGGGCCCCAAACGCATGCTTCCTGTGGTCTAGCCCAGGGGAAGCCCTTCCGTGGGG
GCCCCGGCTTTGAGGGATGCCACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGGTTCGCC
GGGGGCTGCTTGCGTGGATTTCCCGGGTGTGGTTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
CCTGTACATGTTGGCCTGCACCCCCAAAAGGTGACGAGGAGCAGCTGGCACTGCCCAGGGCCAAACAGC
CCCACGGGGAAGGAGGGGTACCAGGCCGTCTTCAGGAGTGGGAGGAGCAGCACCAGCAACTACGTGA
GCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAAGCAGCTCAGGAG
TGGGCAGTACCAAGCCAGCGATGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT
GCCGACCTCCTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGG
CCACAGAGTATGCAGCAGTGCTTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGAGCTGG
CCTTACCCGCCACCCCGAGGAGAAGCCTGTGAGGAAGGACAAGCGGGATGAGTTGGTGGAAAGCCATT
GAATCAGCCTTGAGAGCCCTGAACAATCTGCAGAGAACAGCCCAATCACCCTCCTTACACGGCCT
CTGATTTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTATGAGCTCACCTTCAA
AGGGGACCACAAACACGAATTCAAACGGCTCATCTTATTTTCGACCATTCAGCCCCATCATGAAAGTG
AAAAATGAAAGCTCAACATGGCCAAACAGCTTATCAATGTTATCGTGCCTTAGCAAAAAAGGGTGG
ACAAGTTCCGGCAGTTTCATGCAGAATTTCAAGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCT
CACTGTTGTTTACTTTGGGAAAGAAGAAATAAATGAAGTCAAAGGAATACTTGAAAACACTTCCAAA
GCTGCCAACTTCAGGAACCTTTACCTTCATCCAGCTGAATGGAGAATTTTCTCGGGGAAAGGGACTTG
ATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGTCCTTCTCTTTTCTGTGATGTGGACATCTCACTT
CACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAGGTATTTTATCCAGTT
CTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGGAAACAGC
AGCTGGTCATAAAGAAGGAAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTGTCTAGTATCG
GTACAGACTTCATCAATATAGGTGGGTTTGCATCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCAC
CTTTATCGCAAGTATCTCCACAGCAACCTCATAGTGGTACGGACGCTGTGCGAGGACTCTTCCACC
TCTGGCATGAGAAGCGCTGCATGGACGAGCTGACCCCCGAGCAGTACAAGATGTGCATGCAGTCCAA
GGCCATGAACGAGGCATCCACGGCCAGCTGGGCATGCTGGTGTTCAGGCACGAGATAGAGGCTCAC
TCTTCGAAACAGAAACAGAAAGCAAGTGAACAAAAAACAATGAATCCAGAGAGAGATTTGTGGGAGA
CACTTTTTCTTTCTTTTGTCAATTACTGAAAGTGGCTGCAACAGAGAAAAGACTTCCATAAAGGACG
ACAAAAGAATTGGACTGATGGGTGAGAGATGAGAAAAGCCTCCGATTTCTCTCTGTGGGCTTTTTTAC
AACAGAAATCAAAATCTCCGCTTTGCCTGCAAAAGTAACCCAGTTGCACCCTGTGAAGTGTCTGACA
AAGGCAGAATGCTTGTGAGATTATAAGCCTAATGGTGTGGAGGTTTTTGTATGGTGTGTACAAATCACT
GAGACCTGTTGTTTTGTGTGCTCATTGAAATATTCATGATTTAAGAGCAGTTTTTGTAATAAATTCAT
TAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGAGCCTATCAGCAGGGCTCTAGTTTCTAGG
AATGCTAAAATATCAGAAGGCAGGAGAGGAGATAGGCTTATTATGATACTAGTGAGTACATTAAAGTA
AAATAAAATGGACCAGAAAAGAAAAGAAACCATAAATATCGTGTATATTTTCCCCAAGATTAAACCA
AAAATAATCTGCTTATCTTTTTGGTTGTCCTTTTAACTGTCTCCGTTTTTTTTCTTTTATTTAAAAAT
GCACTTTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAAGAGA
GCACAAGTTGGCCTACATTTTATATTTTAAAGAGATACTTTGAGATGCATTATGAGAAGCTTCA
GTTCAAAGCATCAAAATGATGCCATATCCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAAT
GTCAGGCATTGAGACATAGGGAAGGAATGTTTTGTACTAATACAGACGTACAGATACTTTCTCTGAA
GAGTATTTTCGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGACACTTTCTGCTTTACAGAA
AAGGAAACTCATTACAGACTGGTGATATCGTGATGTACCTAAAAGTCAGAAACCACATTTTCTCCTCA
GAAGTAGGGACCGCTTTCTTACCTGTTTAAATAAACCAGGATATACCGTGTGAACCAACAATCTCT
TTTTCAAAACAGGGTGCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAAATATATATAT
ATATATATATATTGTGAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAAGTTTGTGTACAT
GTTATCCACCCAGGCCAGGTGGAAGTAACTGAATATTTTTTAAATTAAGCAGTTTCTACTCAATCA
CCAAGATGCTTCTGAAAATTGCATTTTATTACCATTTCAAACATTTTTTTTAAAAATAAATACAGTTA
ACATAGAGTGGTTTTCTTCATTTCATGTGAAAATTATTAGCCAGCACCAGATGCATGAGCTAATTATCT
CTTTGAGTCTTTGCTTCTGTTTGTCTCAGTAACTCATTGTTTAAAGCTTCAAGAACATTCAAGC
TGTTGGTGTGTTAAAAAATGCATTGTATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT
AGGCCATGAATGGAAGGTGGTATTGCACAGCTAATAAATATGATTGTGGATATGAA

276/330

FIGURE 276

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQ
EWEEQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFL
HSQVDKAEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPV RKDKRDELVEAIES
ALETLNPAENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSP
MKVKNEKLN MANTLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVH LTVVYFGKEEINEVK
GILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCR
LNTQPGKKVFYPVLFSQYNPGIIYGHHD AVPPLEQQLV IKKETGFWRDFGFGMTCQYRSDFI
NIGGFDL DIKGWGGEDVHLYRKYLHSNLIVVRTPVRGLFHLWHEKRCMDELTP EQYKMCMQS
KAMNEASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT

277/330

FIGURE 277

GAAAGA**ATG**TTGTGGCTGCTCTTTTTTCTGGTGACTGCCATTCATGCTGAACTCTGTCAACC
AGGTGCAGAAAATGCTTTTAAAGTGAGACTTAGTATCAGAACAGCTCTGGGAGATAAAGCAT
ATGCCTGGGATACCAATGAAGAATACCTCTTCAAAGCGATGGTAGCTTTCTCCATGAGAAAA
GTTCCCAACAGAGAAGCAACAGAAATTTCCCATGTCCTACTTTGCAATGTAACCCAGAGGGT
ATCATTCTGGTTTGTGGTTACAGACCCTTCAAAAAATCACACCCTTCCTGCTGTTGAGGTGC
AATCAGCCATAAGAATGAACAAGAACCGGATCAACAATGCCTTCTTTCTAAATGACCAAAC
CTGGAATTTTTTAAAAATCCCTTCCACACTTGCACCACCCATGGACCCATCTGTGCCCATCTG
GATTATTATATTTGGTGTGATATTTTGCATCATCATAGTTGCAATTGCACTACTGATTTTAT
CAGGGATCTGGCAACGTAGAAGAAAGAACAAAGAACCATCTGAAGTGGATGACGCTGAAGAT
AAGTGTGAAAACATGATCACAATTGAAAATGGCATCCCCTCTGATCCCCTGGACATGAAGGG
GGGCATATTAATGATGCCTTCA**TGA**CAGAGGATGAGAGGGCTCACCCCTCTCTGAAGGGCTGT
TGTTCTGCTTCCTCAAGAAATTAAACATTTGTTTCTGTGTGACTGCTGAGCATCCTGAAATA
CCAAGAGCAGATCATATATTTTGTTCACCATTCTTCTTTTGTAATAAATTTTGAATGTGCT
TGAAAGTGAAAAGCAATCAATTATACCCACCAACACCACTGAAATCATAAGCTATTCACGAC
TCAAAATATTCTAAAATATTTTTCTGACAGTATAGTGTATAAATGTGGTCATGTGGTATTTG
TAGTTATTGATTTAAGCATTTTTAGAAATAAGATCAGGCATATGTATATATTTTCACACTTC
AAAGACCTAAGGAAAAATAAATTTTCCAGTGGAGAATACATATAATATGGTGTAGAAATCAT
TGAAAATGGATCCTTTTTTGACGATCACTTATATCACTCTGTATATGACTAAGTAAACAAAAG
TGAGAAGTAATTATTGTAAATGGATGGATAAAAATGGAATTACTCATATACAGGGTGGAATT
TTATCCTGTTATCACACCAACAGTTGATTATATATTTTTCTGAATATCAGCCCCTAATAGGAC
AATTCTATTTGTTGACCATTTCTACAATTTGTAAAAGTCCAATCTGTGCTAACTTAATAAAG
TAATAATCATCTCTTTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

278/330

FIGURE 278

MLWLLFFLVTAIHAELCQPGAENAFKVRLSIRTALGDKAYAWDTNEEYLFKAMVAFSMRKVP
NREATEISHVLLCNVTQRVSFVFWVTDPSKNHTLPAVEVQSAIRMNKNRINNAFFLNDQTL
FLKIPSTLAPPMDPSVPIWIIIFGVIFCIIIVAIALLILSGIWQRRRKNKEPSEVDDAEDKC
ENMITIENGIPSDPLDMKGGILMMPS

279/330

FIGURE 279

AACTCAAACCTCTCTCTGCGGAAACGCGGTGCTTGCTCCTCCCGGAGTGGCCTTGGCAGG
GTGTTGGAGCCCTCGGTCTGCCCCGTCCGGTCTCTGGGGCCAAGGCTGGGTTTCCCTC**ATGT**
ATGGCAAGAGCTCTACTCGTGCGGTGCTTCTTCTCCTTGGCATAACAGCTCACAGCTCTTTGG
CCTATAGCAGCTGTGGAAATTTATACCTCCCGGGTGCTGGAGGCTGTTAATGGGACAGATGC
TCGGTTAAAATGCACTTTCTCCAGCTTTGCCCTGTGGGTGATGCTCTAACAGTGACCTGGA
ATTTTCGTCCTCTAGACGGGGGACCTGAGCAGTTTGTATTCTACTACCACATAGATCCCTTC
CAACCCATGAGTGGGCGGTTTAAGGACCGGGTGTCTTGGGATGGGAATCCTGAGCGGTACGA
TGCCTCCATCCTTCTCTGGAAACTGCAGTTCGACGACAATGGGACATACACCTGCCAGGTGA
AGAACCCACCTGATGTTGATGGGGTGATAGGGGAGATCCGGCTCAGCGTCGTGCACACTGTA
CGCTTCTCTGAGATCCACTTCCTGGCTCTGGCCATTGGCTCTGCCTGTGCACTGATGATCAT
AATAGTAATTGTAGTGGTCCTCTTCCAGCATTACCGGAAAAAGCGATGGGCCGAAAGAGCTC
ATAAAGTGGTGGAGATAAAATCAAAGAAGAGGAAAGGCTCAACCAAGAGAAAAAGGTCTCT
GTTTATTTAGAAAGACACAGAC**TAA**CAATTTTAGATGGAAGCTGAGATGATTTCCAAGAACAA
GAACCCTAGTATTTCTTGAAGTTAATGGAACTTTTCTTTGGCTTTTCCAGTTGTGACCCGT
TTTCCAACCAGTTCTGCAGCATATTAGATTCTAGACAAGCAACACCCCTCTGGAGCCAGCAC
AGTGCTCCTCCATATCACAGTCATACACAGCCTCATTATTAAGGTCTTATTTAATTTTCA
GTGTAAATTTTTTCAAGTGCTCATTAGGTTTTATAACAAGAAGCTACATTTTTTGGCCTTAA
GACACTACTTACAGTGTTATGACTTGTATACACATATATTGGTATCAAAGGGGATAAAAGCC
AATTTGTCTGTTACATTTCTTTTACGTATTTCTTTTAGCAGCACTTCTGCTACTAAAGTTA
ATGTGTTTACTCTCTTTCTTCCACATTCTCAATTAAAAGGTGAGCTAAGCCTCCTCGGTG
TTTCTGATTAACAGTAAATCCTAAATTCAAAGTGTAAATGACATTTTTATTTTTATGTCTC
TCCTTAACTATGAGACACATCTTGTTTTACTGAATTTCTTTCAATATTCCAGGTGATAGATT
TTTGTCG

280/330

FIGURE 280

MYGKSSTRAVLLLLGIQLTALWPAAVEIYTSRVLEAVNGTDARLKCTFSSFAPVGDALTVT
WNFRPLDGGPEQFVFYYHIDPFQPMSGGRFKDRVSWDGNPERYDASILLWKLQFDDNGTYTCQ
VKNPPDVDGVIGEIRLSVVHTVRFSEIHFLALAIGSACALMIIIVIVVVL FQHYRKKRWAER
AHKVVEIKSKEEERLNQEKVSVYLEDTD

281/330

FIGURE 281

GCATTTTTGTCTGTGCTCCCTGATCTTCAGGTCACCACC**ATGA**AAGTTCTTAGCAGTCCTGGT
ACTCTTGGGAGTTTCCATCTTTCTGGTCTCTGCCCAGAATCCGACAACAGCTGCTCCAGCTG
ACACGTATCCAGCTACTGGTCCTGCTGATGATGAAGCCCCTGATGCTGAAACCACTGCTGCT
GCAACCACTGCGACCACTGCTGCTCCTACCACTGCAACCACCGCTGCTTCTACCACTGCTCG
TAAAGACATTCCAGTTTTACCCAAATGGGTTGGGGATCTCCCGAATGGTAGAGTGTGTCCCT**T**
GAGATGGAATCAGCTTGAGTCTTCTGCAATTGGTCACAACCTATTCATGCTTCCTGTGATTTC
ATCCAACCTACTTACCTTGCCTACGATATCCCCTTTATCTCTAATCAGTTTATTTTCTTTCAA
ATAAAAAATAACTATGAGCAACATAAAAAAAAAAAAAA

282/330

FIGURE 282

MKFLAVLVLLGVSI FLVSAQNPTTAAPADTYPATGPADDEAPDAETTAAATTATTAAPTAT
TAASTTARKDIPVLPKWVGDL PNGRVCP

283/330

FIGURE 283

GGACTCTGAAGGTCCCAAGCAGCTGCTGAGGCCCCCAAGGAAGTGGTTCCAACCTTGGACCC
CTAGGGGTCTGGATTTGCTGGTTAACAAGATAACCTGAGGGCAGGACCCCATAGGGGA**ATGC**
TACCTCCTGCCCTTCCACCTGCCCTGGTGTTACGGGTGGCCTGGTCCCTCCTTGCCGAGAGA
GTGTCCTGGGTCAGGGACGCAGAGGACGCTCACAGACTCCAGCCCTTTGTTACCGAGAGGAC
ACTTGGCAAGGTCCAGCGATGGTCCGGAGTCCACACACAGACTGGCGGCAGGGCAGGAGGGG
GACAGTTCTGTTGTGCTTGGTTGGACAGTAAGAGGGTCTTGCCAGTCCAGGGTGGGGGGCG
GCAAACCTCCATAAAGAACCAGAGGGTCTGGGCCCCGGCCACAGAGTCATCTGCCCAGCTCCT
CTGCTGCTGGCCAGTGGGAGTGGCACGAGGTGGGGCTTTGTGCCAG**TAA**AACCACAGGCTGG
ATTTGCCTGCGGGCCATGGTCCCTGTCTAGGGCAGCAATTCTCAACCTTCTTGCTCTCAGGA
CCCCAAAGAGCTTTCATTGTATCTATTGATTTTTTACCACATTAGCAATTAAACTGAGAAAT
GGGCCGGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGAT
CACCTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTACTAAAAA
TACAAAAAATTAGCCAGGCACAGTGGTGTGCACTGGTAGTCCCAGTTACTCGGGAGGCTGAG
GCAGGAAAATCGCTTGAACCCAGGAGGCGGACGTTGCGGTGAGCCGAGATCGCGCCGCTGAT
TCCAGCCTGGGCGACAAGAGTGAGACTCCATCTCACACA

284/330

FIGURE 284

MLPPALPPALVFTVAWSLLAERVSWVRDAEDAHRLQPFVTERTLGKVQRWSGVHTQTGGRAG
GGQFCCAWLDSKRVLASPGWGAANSIKNQRVWAPATESSAQLCCWPVGVARGGALCQ

285/330

FIGURE 285

GTC**ATG**CCAGTGCCTGCTCTGTGCCTGCTCTGGGCCCTGGCAATGGTGACCCGGCCTGCCTCA
GCGGCCCCCATGGGCGGCCCAGAACTGGCACAGCATGAGGAGCTGACCCTGCTCTTCCATGG
GACCCTGCAGCTGGGCCAGGCCCTCAACGGTGTGTACAGGACCACGGAGGGACGGCTGACAA
AGGCCAGGAACAGCCTGGGTCTCTATGGCCGCACAATAGAACTCCTGGGGCAGGAGGTCAGC
CGGGGCCGGGATGCAGCCCAGGAACTTCGGGCAAGCCTGTTGGAGACTCAGATGGAGGAGGA
TATTCTGCAGCTGCAGGCAGAGGCCACAGCTGAGGTGCTGGGGGAGGTGGCCCAGGCACAGA
AGGTGCTACGGGACAGCGTGCAGCGGCTAGAAGTCCAGCTGAGGAGCGCCTGGCTGGGCCCT
GCCTACCGAGAATTTGAGGTCTTAAAGGCTCACGCTGACAAGCAGAGCCACATCCTATGGGC
CCTCACAGGCCACGTGCAGCGGCAGAGGCGGGAGATGGTGGCACAGCAGCATCGGCTGCGAC
AGATCCAGGAGAGACTCCACACAGCGGCGCTCCCAGCC**TGA**ATCTGCCTGGATGGAAGTGA
GACCAATCATGCTGCAAGGAACACTTCCACGCCCCGTGAGGCCCTGTGCAGGGAGGAGCTG
CCTGTTCACCTGGGATCAGCCAGGGCGCCGGGCCCCACTTCTGAGCACAGAGCAGAGACAGAC
GCAGGCGGGGACAAAGGCAGAGGATGTAGCCCCATTGGGGAGGGGTGGAGGAAGGACATGTA
CCCTTTCATGCCTACACACCCCTCATTAAGCAGAGTCGTGGCATTTCAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

286/330

FIGURE 286

MPVPALCLLWALAMVTRPASAAPMGGPELAQHEELTLLFHGTLQLGQALNGVYRTTEGRLTK
ARNSLGLYGRTIELLGQEVSRGRDAAQELRASLLETQMEEDILQLQAEATAEVLGEVAQAQK
VLRDSVQRLEVQLRSAWLGPAYREFEVLKAHADKQSHILWALTGHVQRQRREMVAQQHRLRQ
IQERLHTAALPA

287/330

FIGURE 287

GGCAAC**ATG**GCTCAGCAGGCTTGCCCCAGAGCCATGGCAAAGAATGGACTTGTAATTTGCAT
CCTGGTGATCACCTTACTCCTGGACCAGACCACCAGCCACACATCCAGATTAAAAGCCAGGA
AGCACAGCAAACGTCGAGTGAGAGACAAGGATGGAGATCTGAAGACTCAAATTGAAAAGCTC
TGGACAGAAGTCAATGCCTTGAAGGAAATTCAAGCCCTGCAGACAGTCTGTCTCCGAGGCAC
TAAAGTTCACAAGAAATGCTACCTTGCTTCAGAAGGTTTGAAGCATTTCCATGAGGCCAATG
AAGACTGCATTTCCAAAGGAGGAATCCTGGTTATCCCCAGGAACTCCGACGAAATCAACGCC
CTCCAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGACTTTTGGCTGGGCATCAATGA
CATGGTCACGGAAGGCAAGTTTGTGACGTCAACGGAATCGCTATCTCCTTCCTCAACTGGG
ACCGTGCACAGCCTAACGGTGGCAAGCGAGAAAACGTGTCTCCTGTTCTCCAATCAGCTCAG
GGCAAGTGGAGTGATGAGGCCTGTGCGCAGCAGCAAGAGATACATATGCGAGTTCACCATCCC
TAA**TAG**GTCTTTCTCCAATGTGTCTCCTCCAAGCAAGATTCATCATAACTTATAGGTTTCATGA
TCTCTAAGATCAAGTAAAAATCATAATTTTTACTTATTAAAAAATTGCAACACAAGATCAAT
GTCCATAGCAATATGATAGCATCAGCCAATTTTGCTAACACATTTCTTTGGGATTTTGCCCT
TCCTGGGGTATAGGGGATCAGAAATATTGATCCATGTGCACGCAGATAAAATGGCTTCTGCT
AAACAGACTAAAATCTTTCTCTCTAGTCTTTCTCACTTGTACAAACCCAGTTTGTGTTTCAA
AAATCACAGTAGCAATGCAACTCATCACTCTAGAAAAGCAAGCTTAGGCTACCTGAAAGATT
TTCCCTTGGAAGTTTAGCGTATGTTTGACTAACAAAAATTCCTTACATCAGAGACTCTAGGT
GCTATATAATCCAAAACTTTTCAGCCTGTTGCTCATTCTGTCCCATGCTGGCAATAATACC
TTGTCAGCCCATTACCCTTATTTTGAATTGCTCCATCTCCTGGTGGGACTTGTATCTTGTCT
GCCATATCAGAACACAAACCCCTGAAGAGGTTCTGATTTGATTTTTTTTTTTTTCTTCATGCC
TACCCTTTTTTTTGAAGTTTCCAGCCGCAATTTGAAATGAAATGACAAGGTGTATATTTGAT
CAATTTTCATTCCCACCATTGCATTACAACCTCTAACTTAAATGGGTAACCTAAGGCATAT
CAAAGAAGCAGATTGCATGATAAACGGAAATAGAAAAAAGAACCTACATTTATTTTGCTTT
AGCATCCTTACTCTCACCTTTTATGAGATTGAGAGTGGACTTACATTTCTTTTTTACATTT
TCGTATATTTATTTTTTTTAGCCATCATTATATGTTTAAGTCTATTATGGGCAACCAATCTT
TGGAAGCTGAAAACCTGAATTTAAAGAATGCTATCTTGGAATTTGCATACGTCTGTGCAATT
TTTTATTCTGCCTAGTGCTATTCTGCTTGTTTAACTAGATTGTACAAAATAACTTCATTGCT
TAATATCAAATTACAAAGTTTAGACTTGGAGGGAAATGGGCTTTTTTAGAAGCAAACAATTTT
AAATATATTTTGTCTTCAAATAAATAGTGTTTAAACATTGAATGTGTTTTGTGAACAATAT
CCCACTTTGCAAACTTTAACTACACATGCTTGGAATTAAGTTTTAGCTGTTTTCATTGCTCA
ATAATAAAGCCTGAATTCTGATCAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

288/330

FIGURE 288

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHSKRRVRDKDGD LKTQIEKLWT
EVNALKEIQALQTVCLRGTKVHKKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQ
DYGKRSLPGVNDFWLGINDMVTEGKFVDVNGIAISFLNWDRAQPNGGKRENCVLFSQSAQ GK
WSDEACRSSKRYICEFTIPK

289/330

FIGURE 289

GCGAGGACCGGGTATAAGAAGCCTCGTGCCCTTGCCCGGGCAGCCGCAGGTTCCCCGCGCGC
CCCGAGCCCCCGCGCC**ATGA**AAGCTCGCCGCCCTCCTGGGGCTCTGCGTGGCCCTGTCCTGCA
GCTCCGCTGCTGCTTTCTTAGTGGGCTCGGCCAAGCCTGTGGCCCAGCCTGTCGCTGCGCTG
GAGTCGGCGGCGGAGGCCGGGGCCGGGACCCTGGCCAACCCCTCGGCACCCTCAACCCGCT
GAAGCTCCTGCTGAGCAGCCTGGGCATCCCCGTGAACCACCTCATAGAGGGCTCCCAGAAGT
GTGTGGCTGAGCTGGGTCCCCAGGCCGTGGGGGCCGTGAAGGCCCTGAAGGCCCTGCTGGGG
GCCCTGACAGTGTTTGGC**TGA**AGCCGAGACTGGAGCATCTACACCTGAGGACAAGACGCTGCC
CACCCGCGAGGGCTGAAAACCCCGCCGCGGGGAGGACCGTCCATCCCCTTCCCCCGGCCCT
CTCAATAAACGTGGTTAAGAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAA

290/330

FIGURE 290

MKLAALLGLCVALSCSSAAAFVLGSAKPVAQPVAALESAAEAGAGTLANPLGTINPLKLLLS
SLGIPVNHLEGSQKCVaelGPQAVGAVKALKALLGALTVFG

291/330

FIGURE 291

TGAAGGACTTTTCCAGGACCCCAAGGCCACACACTGGAAGTCTTGCAGCTGAAGGGAGGCACT
CCTTGGCCTCCGCAGCCGATCAC**ATGA**AAGGTGGTGCCAAGTCTCCTGCTCTCCGTCCTCCTG
GCACAGGTGTGGCTGGTACCCGGCTTGGCCCCCAGTCCTCAGTCGCCAGAGACCCAGCCCC
TCAGAACCAGACCAGCAGGGTAGTGCAGGCTCCCAGGGAGGAAGAGGAAGATGAGCAGGAGG
CCAGCGAGGAGAAGGCCGGTGAGGAAGAGAAAGCCTGGCTGATGGCCAGCAGGCAGCAGCTT
GCCAAGGAGACTTCAAACCTTCGATTTCAGCCTGCTGCGAAAGATCTCCATGAGGCACGATGG
CAACATGGTCTTCTCTCCATTTGGCATGTCCTTGGCCATGACAGGCTTGATGCTGGGGGCCA
CAGGGCCGACTGAAACCCAGATCAAGAGAGGGCTCCACTTGCAGGCCCTGAAGCCCACCAAG
CCCGGGCTCCTGCCTTCCCTCTTTAAGGGACTCAGAGAGACCCTCTCCCGCAACCTGGAACCT
GGGCCTCTCACAGGGGAGTTTTGCCTTCATCCACAAGGATTTTGATGTCAAAGAGACTTTCT
TCAATTTATCCAAGAGGTATTTTGATACAGAGTGCCTGCTGCTATGAATTTTCGCAATGCCTCA
CAGGCCAAAAGGCTCATGAATCATTACATTAACAAAGAGACTCGGGGGAAAATTTCCAAACT
GTTTGATGAGATTAATCCTGAAACCAAATTAATTCTTGTGGATTACATCTTGTTCAAAGGGA
AATGGTTGACCCCATTTGACCCTGTCTTCACCGAAGTCGACACTTTCCACCTGGACAAGTAC
AAGACCATTAAGGTGCCCATGATGTACGGTGCAGGCAAGTTTGCCTCCACCTTTGACAAGAA
TTTTCGTTGTTCATGTCCTCAAACCTGCCCTACCAAGGAAATGCCACCATGCTGGTGGTCTCA
TGGAGAAAATGGGTGACCACCTCGCCCTTGAAGACTACCTGACCACAGACTTGGTGGAGACA
TGGCTCAGAAACATGAAAACAGAAACATGGAAGTTTTCTTTCCGAAGTTCAAGCTAGATCA
GAAGTATGAGATGCATGAGCTGCTTAGGCAGATGGGAATCAGAAGAATCTTCTCACCCCTTG
CTGACCTTAGTGAACCTCTCAGCTACTGGAAGAAATCTCCAAGTATCCAGGGTTTTACGAAGA
ACAGTGATTGAAGTTGATGAAAGGGGCACTGAGGCAGTGGCAGGAATCTTGTCAGAAATTAC
TGCTTATTCCATGCCTCCTGTTCATCAAAGTGGACCGGCCATTTTCATTTTCATGATCTATGAAG
AAACCTCTGGAATGCTTCTGTTTTCTGGGCAGGGTGGTGAATCCGACTCTCCTAT**TAA**ATTCAGG
ACATGCATAAGCACTTCGTGCTGTAGTAGATGCTGAATCTGAGGTATCAAACACACACAGGA
TACCAGCAATGGATGGCAGGGGAGAGTGTTCCTTTTGTTCTTAAGTATAGTTTAGGGTGTTCTC
AAATAAATACAGTAGTCCCCACTTATCTGAGGGGGATACATTCAAAGACCCCCAGCAGATGC
CTGAAACGGTGGACAGTGCTGAACCTTATATATATTTTTTCTTACACATACATACCTATGAT
AAAGTTTAATTTATAAATTAGGCACAGTAAGAGATTAACAATAATAACAACATTAAGTAAAA
TGAGTTACTTGAACGCAAGCACTGCAATACCATAACAGTCAAACCTGATTATAGAGAAGGCTA
CTAAGTGACTCATGGGCGAGGAGCATAGACAGTGTGGAGACATTGGGCAAGGGGAGAATTCA
CATCCTGGGTGGGACAGAGCAGGACGATGCAAGATTCCATCCCACTACTCAGAATGGCATGC
TGCTTAAGACTTTTAGATTGTTTTATTTCTGGAATTTTTTCATTTAATGTTTTTGGACCATGGT
TGACCATGGTTAACTGAGACTGCAGAAAGCAAAACCATGGATAAGGGAGGACTACTACAAAA
GCATTAAATTGATACATATTTTTTAAAAA

292/330

FIGURE 292

MKVVP S L L L S V L L A Q V W L V P G L A P S P Q S P E T P A P Q N Q T S R V V Q A P R E E E E D E Q E A S E E K A G E
E E K A W L M A S R Q Q L A K E T S N F G F S L L R K I S M R H D G N M V F S P F G M S L A M T G L M L G A T G P T E T Q I
K R G L H L Q A L K P T K P G L L P S L F K G L R E T L S R N L E L G L S Q G S F A F I H K D F D V K E T F F N L S K R Y F
D T E C V P M N F R N A S Q A K R L M N H Y I N K E T R G K I P K L F D E I N P E T K L I L V D Y I L F K G K W L T P F D P
V F T E V D T F H L D K Y K T I K V P M M Y G A G K F A S T F D K N F R C H V L K L P Y Q G N A T M L V V L M E K M G D H L
A L E D Y L T T D L V E T W L R N M K T R N M E V F F P K F K L D Q K Y E M H E L L R Q M G I R R I F S P F A D L S E L S A
T G R N L Q V S R V L R R T V I E V D E R G T E A V A G I L S E I T A Y S M P P V I K V D R P F H F M I Y E E T S G M L L F
L G R V V N P T L L

293/330

FIGURE 293

CTGGGATCAGCCACTGCAGCTCCCTGAGCACTCTCTACAGAGACGCGGACCCCAGAC**ATG**AG
GAGGCTCCTCCTGGTCACCAGCCTGGTGGTTGTGCTGCTGTGGGAGGCAGGTGCAGTCCCAG
CACCCAAGGTCCCTATCAAGATGCAAGTCAAACACTGGCCCTCAGAGCAGGACCCAGAGAAG
GCCTGGGGCGCCCGTGTGGTGGAGCCTCCGGAGAAGGACGACCAGCTGGTGGTGTCTGTTCCC
TGTCCAGAAGCCGAAACTCTTGACCACCGAGGAGAAGCCACGAGGTCAGGGCAGGGGCCCCA
TCCTTCCAGGCACCAAGGCCTGGATGGAGACCGAGGACACCCTGGGCCGTGTCCTGAGTCCC
GAGCCCGACCATGACAGCCTGTACCACCCTCCGCCTGAGGAGGACCAGGGCGAGGAGAGGCC
CCGGTTGTGGGTGATGCCAAATCACCAGGTGCTCCTGGGACCGGAGGAAGACCAAGACCACA
TCTACCACCCCCAG**TAG**GGCTCCAGGGGCCATCACTGCCCCCGCCCTGTCCCAAGGCCCAGG
CTGTTGGGACTGGGACCCTCCCTACCCTGCCCCAGCTAGACAAATAAACCCCAGCAGGCAAA
AAAAAAAAAAAAAAAAA

294/330

FIGURE 294

MRRLLLVTSLVVVLLWEAGAVPAPKVPIKMQVKHWPSEQDPEKAWGARVVEPPEKDDQLVVL
FPVQKPKLLTTEEKPRGQGRGPILPGTKAWMETEDTLGRVLSPEPDHDSLYHPPPEEDQGEE
RPRLWVMPNHQVLLGPEEDQDHIYHPQ

295/330

FIGURE 295

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTG
TACCCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACA**ATGA**ACCAACTCAGCTTCCTGC
TGTTTCTCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAA
TGGACCTGTTCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCC
TAGTGCATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTTATCTACCAGACCTTCT
GTGACATGACCTCTGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTG CATGAGAATGACATG
CGTGGGAAGTGCACGGTGGGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCC
AGAGGGGGACGGCAACTGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCG
ATGACTACAAGAACCCTGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTG
CCCAATAAGTCCCCCATGCAGCACTGGAGAAACAGCTCCCTGCTGAGGTACCGCACGGACAC
TGGCTTCCTCCAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCAGTGAAAT
ATGGAGAAGGAAAGTGTTGGACTGACAACGGCCCCGGTGATCCCTGTGGTCTATGATTTTGGC
GACGCCCAGAAAACAGCATCTTATTACTCACCTATGGCCAGCGGGAATTCCTGCGGGATT
TGTT CAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAACGCCTTGTGTGCTGGAATGAGGG
TCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAGGATACTTTCCAGAGGCCAGT
CCCCAGCAGTGTGGAGATTTTTCTGGTTTTGATTGGAGTGGATATGGAACCTCATGTTGGTTA
CAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCGT**TGA**GAGTTTTGTG
GGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAACAACCTTACCCA
GTAGCTAGAATGTTAATGGCAGAAGAGAAAAACAATAAATCATATTGACTCAAGAAAAAAA

296/330

FIGURE 296

MNQLSFLLEFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTEN
GVIYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFG
SAEAATSDDYKNPGYYDIQAKDLGIWHVFNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGLI
YQKYPVKYGEKGCWTDNGPVI PVVYDFGDAQKTASYYSPTYGQREFTAGFVQFRVFNNERAAN
ALCAGMRVTGCNTEHHCIGGGGYFPEAS PQQCGDFSGFDWSGYGTHVGYSSSREITEAAVLL
FYR

297/330

FIGURE 297

GCGGAGCCGGCGCCGGCTGCGCAGAGGAGCCGCTCTCGCCGCCGCCACCTCGGCTGGGAGCC
CACGAGGCTGCCGCATCCTGCCCTCGGAACA**ATG**GGACTCGGCGCGCGAGGTGCTTGGGCGG
CGCTGCTCCTGGGGACGCTGCAGGTGCTAGCGCTGCTGGGGGCCGCCCATGAAAGCGCAGCC
ATGGCGGCATCTGCAAACATAGAGAATTCTGGGCTTCCACACAACCTCCAGTGCTAACTCAAC
AGAGACTCTCCAACATGTGCCTTCTGACCATACAAATGAACTTCCAACAGTACTGTGAAAC
CACCAACTTCAGTTGCCTCAGACTCCAGTAATACAACGGTCACCACCATGAAACCTACAGCG
GCATCTAATAACAACACCAGGGATGGTCTCAACAAATATGACTTCTACCACCTTAAAGTC
TACACCCAAAACAACAAGTGTTTCACAGAACACATCTCAGATATCAACATCCACAATGACCG
TAACCCACAATAGTTCAGTGACATCTGCTGCTTCATCAGTAACAATCACAACAACCTATGCAT
TCTGAAGCAAAGAAAGGATCAAAATTTGATACTGGGAGCTTTGTTGGTGGTATTGTATTAAC
GCTGGGAGTTTTATCTATTCTTTACATTGGATGCAAATGTATTACTCAAGAAGAGGCATTC
GGTATCGAACCATAGATGAACATGATGCCATCATT**TAA**GGAAATCCATGGACCAAGGATGGA
ATACAGATTGATGCTGCCCTATCAATTAATTTTGGTTTATTAATAGTTTAAAACAATATTCT
CTTTTTGAAAATAGTATAAACAGGCCATGCATATAATGTACAGTGTATTACGTAAATATGTA
AAGATTCTTCAAGGTAACAAGGGTTTGGGTTTTGAAATAAACATCTGGATCTTATAGACCGT
TCATACAATGGTTTTAGCAAGTTCATAGTAAGACAAACAAGTCCTATCTTTTTTTTTTTGGCT
GGGGTGGGGGCATTGGTCACATATGACCAGTAATTGAAAGACGTCATCACTGAAAGACAGAA
TGCCATCTGGGCATACAAATAAGAAGTTTGTCACAGCACTCAGGATTTTGGGTATCTTTTGT
AGCTCACATAAAGAACTTCAGTGCTTTTCAGAGCTGGATATATCTTAATTACTAATGCCACA
CAGAAATTATACAATCAAACCTAGATCTGAAGCATAATTTAAGAAAAACATCAACATTTTTTG
TGCTTTAACTGTAGTAGTTGGTCTAGAAACAAAATACTCC

298/330

FIGURE 298

MGLGARGAWAALLLGTLQVLALLGAAHESAAMAASANIENSGLPHNSSANSTETLQHVP
SDH
TNETSNSTVKPPTSVASDSSNTTVTTMKPTAASNTTTPGMVSTNMTSTTLKSTPKTTSV
SQN
TSQISTSTMTVTHNSSVTSAASSVTITTTMHSEAKKGSKFDTGSFVGGIVLTLGVLSI
LYIG
CKMYYSRRGIRYRTIDEHDAII

299/330

FIGURE 299

CAGCCGGGTCCCAAGCCTGTGCCTGAGCCTGAGCCTGAGCCTGAGCCCGAGCCGGGAGCCGG
TCGCGGGGGCTCCGGGCTGTGGGACCGCTGGGCCCCCAGCG**ATG**GCGACCCTGTGGGGAGGC
CTTCTTCGGCTTGGCTCCTTGCTCAGCCTGTCGTGCCTGGCGCTTCCGTGCTGCTGCTGGC
GCAGCTGTCAGACGCCGCCAAGAATTCGAGGATGTCAGATGTAAATGTATCTGCCCTCCCT
ATAAAGAAAATTCTGGGCATATTTATAATAAGAACATATCTCAGAAAGATTGTGATTGCCTT
CATGTTGTGGAGCCCATGCCTGTGCGGGGGCCTGATGTAGAAGCATACTGTCTACGCTGTGA
ATGCAAATATGAAGAAAGAAGCTCTGTCACAATCAAGGTTACCATTATAATTTATCTCTCCA
TTTTGGGCCTTCTACTTCTGTACATGGTATATCTTACTCTGGTTGAGCCCATACTGAAGAGG
CGCCTCTTTGGACATGCACAGTTGATACAGAGTGATGATGATATTGGGGATCACCAGCCTTT
TGCAAATGCACACGATGTGCTAGCCCGCTCCCGCAGTCGAGCCAACGTGCTGAACAAGGTAG
AATATGCACAGCAGCGCTGGAAGCTTCAAGTCCAAGAGCAGCGAAAGTCTGTCTTTGACCGG
CATGTTGTCCTCAGC**TAA**TTGGGAATTGAATTCAAGGTGACTAGAAAGAAACAGGCAGACAA
CTGGAAAGAACTGACTGGGTTTTGCTGGGTTTCATTTTAATACCTTGTTGATTTACCAACT
GTTGCTGGAAGATTCAAACTGGAAGCAAAACTTGCTTGATTTTTTTTTCTTGTTAACGTA
ATAATAGAGACATTTTTTAAAGCACACAGCTCAAAGTCAGCCAATAAGTCTTTTCCTATTTG
TGACTTTTACTAATAAAAATAAATCTGCCTGTAAATTATCTTGAAGTCCTTTACCTGGAACA
AGCACTCTCTTTTTTACCACATAGTTTTTAACTTGACTTTCAAGATAATTTTCAGGGTTTTTG
TTGTTGTTGTTTTTTGTTTGTTTGTTTGGTGGGAGAGGGGAGGGATGCCTGGGAAGTGGTT
AACAACTTTTTTCAAGTCACTTTACTAAACAACTTTTGTAATAGACCTTACCTTCTATTT
TCGAGTTTCATTTATATTTTGCAGTG TAGCCAGCCTCATCAAAGAGCTGACTTACTCATTTG
ACTTTTGCCTGACTGTATTATCTGGGTATCTGCTGTGTCTGCACTTCATGGTAAACGGGAT
CTAAAATGCCTGGTGGCTTTTCACAAAAGCAGATTTTCTTCATGTACTGTGATGTCTGATG
CAATGCATCCTAGAACAACTGGCCATTTGCTAGTTTACTCTAAAGACTAAACATAGTCTTG
GTGTGTGTGGTCTTACTCATCTTCTAGTACCTTTAAGGACAAATCCTAAGGACTTGGACACT
TGCAATAAAGAAATTTTATTTTAAACCCAAGCCTCCCTGGATTGATAATATATACACATTTG
TCAGCATTTCCGGTCGTGGTGAGAGGCAGCTGTTTGAGCTCCAATATGTGCAGCTTTGAACT
AGGGCTGGGGTTGTGGGTGCCTCTTCTGAAAGGTCTAACCATTATTGGATAACTGGCTTTTT
TCTTCCTATGTCCTCTTTGGAATGTAACAATAAAAATAATTTTTGAAACATCAA

300/330

FIGURE 300

MATLWGGLRLGSLLSLSCLALSVLLLAQLSDAAKNFEDVRCKCICPPYKENS
GHYINKNIS
QKDCDCLHVVEPMPVRGPDVEAYCLRCECKYEERSVTIKVTIIYLSILG
LLLLYMYLTL
VEPILKRRLFGHAQLIQSDDDIGDHQPFANAHDVLRSSRSRANVLNKVEYA
QQRWKLQVQEQ
RKSVFDRHVLS

301/330

FIGURE 301

GCACCTGCGACCACCGTGAGCAGTC**ATG**GCGTACTCCACAGTGCAGAGAGTCGCTCTGGCTT
CTGGGCTTGTCTGGCTCTGTCGCTGCTGCTGCCCCAAGGCCTTCCTGTCCCGCGGGAAGCGG
CAGGAGCCGCGCCGACACCTGAAGGAAAATTGGGCCGATTTCCACCTATGATGCATCATCA
CCAGGCACCCTCAGATGGCCAGACTCCTGGGGCTCGTTTCCAGAGGTCTCACCTTGCCGAGG
CATTTGCAAAGGCCAAAGGATCAGGTGGAGGTGCTGGAGGAGGAGGTAGTGGAAGAGGTCTG
ATGGGGCAGATTATTCCAATCTACGGTTTTGGGATTTTTTTATATATACTGTACATTCTATT
TAAGGTAAGTAGAATCATCCTAATCATATTACATCAAT**TGA**AAATCTAATATGGCGATAAAAA
TCATTGTCTACATTAAAACTTCTTATAGTTCATAAAATTATTTCAAATCCATCATCTCTTTA
AATCCTGCCTCCTCTTCATGAGGTACTTAGGATAGCCATTATTTTCAGTTTTCACATAAGAATG
TTTACTCAATGTTTAAAGTGTTTTGCCCCAAAATTCACAATAACAAGGCAGAACTAGGACTT
GAACATGGATCTTTTGGTTCTTAATCCAGTGAGTGATACAATTCAATGCACTCCCCTGCCA

302/330

FIGURE 302

MAYSTVQRVALASGLVLALSLLLPKAFLSRGKRQEPPPTPEGKLGRFPMMHHHQAPSDGQT
PGARFQRSHLAEAFKAKGSGGGAGGGGSGRGLMGQIIPYGFIFLYILYILFKVSRIILI
ILHQ

303/330

FIGURE 303

CGGCTCGAGTGCAGCTGTGGGGAGATTTTCAGTGCATTGCCTCCCCTGGGTGCTCTTCATCTT
GGATTTGAAAGTTGAGAGCAGC**ATG**TTTTGCCCACTGAAACTCATCCTGCTGCCAGTGTTAC
TGGATTATTCCTTGGGCCTGAATGACTTGAATGTTTCCCCGCCTGAGCTAACAGTCCATGTG
GGTGATTTCAGCTCTGATGGGATGTGTTTTCCAGAGCACAGAAGACAAATGTATATTCAAGAT
AGACTGGACTCTGTCAACCAGGAGAGCACGCCAAGGACGAATATGTGCTATACTATTACTCCA
ATCTCAGTGTGCCTATTGGGCGCTTCCAGAACCGCGTACACTTGATGGGGGACATCTTATGC
AATGATGGCTCTCTCCTGCTCCAAGATGTGCAAGAGGCTGACCAGGGAACCTATATCTGTGA
AATCCGCCTCAAAGGGGAGAGCCAGGTGTTCAAGAAGGCGGTGGTACTGCATGTGCTTCCAG
AGGAGCCCCAAAGAGCTCATGGTCCATGTGGGTGGATTGATTTCAGATGGGATGTGTTTTCCAG
AGCACAGAAGTGAAACACGTGACCAAGGTAGAATGGATATTTTCAGGACGGCGCGCAAAGGA
GGAGATTGTATTTTCGTTACTACCACAACTCAGGATGTCTGTGGAGTACTCCCAGAGCTGGG
GCCACTTCCAGAATCGTGTGAACCTGGTGGGGGACATTTTCCGCAATGACGGTTCCATCATG
CTTCAAGGAGTGAGGGAGTCAGATGGAGGAACTACACCTGCAGTATCCACCTAGGGAACCT
GGTGTTCAGAAAACCATTTGTGCTGCATGTCAGCCCGGAAGAGCCTCGAACACTGGTGACCC
CGGCAGCCCTGAGGCCTCTGGTCTTGGGTGGTAATCAGTTGGTGATCATTGTGGGAATTGTC
TGTGCCACAATCCTGCTGCTCCCTGTTCTGATATTGATCGTGAAGAAGACCTGTGGAAATAA
GAGTTCAGTGAATTCTACAGTCTTGGTGAAGAACACGAAGAAGACTAATCCAGAGATAAAAG
AAAAACCCTGCCATTTTGAAAGATGTGAAGGGGAGAAACACATTTACTCCCCAATAATTGTA
CGGGAGGTGATCGAGGAAGAAGAACCAAGTGAAAAATCAGAGGCCACCTACATGACCATGCA
CCCAGTTTGGCCTTCTCTGAGGTCAGATCGGAACAACCTCACTTGAAAAAAGTCAGGTGGGG
GAATGCCAAAAACACAGCAAGCCTTT**TGA**GAGAAGATGGAGAGTCCCTTCATCTCAGCAGCGG
TGGAGACTCTCTCCTGTGTGTGTCCTGGGCCACTCTACCAGTGATTTTCAGACTCCCGCTCTC
CCAGCTGTCCTCCTGTCTCATTGTTTGGTCAATACACTGAAGATGGAGAATTTGGAGCCTGG
CAGAGAGACTGGACAGCTCTGGAGGAACAGGCCTGCTGAGGGGAGGGGAGCATGGACTTGGC
CTCTGGAGTGGGACACTGGCCCTGGGAACCAGGCTGAGCTGAGTGGCCTCAAACCCCCCGTT
GGATCAGACCCTCCTGTGGGCAGGGTTCTTAGTGGATGAGTTACTGGGAAGAATCAGAGATA
AAAACCAACCCAAATCAA

304/330

FIGURE 304

MFCPLKLILLPVLLDYSLGLNDLNVSPPELTVHVGDSALMGCVFQSTEDKCIFKIDWTLSPG
EHAKDEYVLYYYSNLSVPIGRFQNRVHLMGDILCNDGSLLLQDVQEADQGTYICEIRLKGES
QVFKKAVVLHVLPEEPKELMVHVGGLIQMGCVFQSTEVKHVTKVEWIFSGRRAKEEIVFRYY
HKL RMSVEYSQSWG HFQNRVNLVGDI FRNDGSIMLQGVRES DGGNYTCSIHLGNLVFKKTIV
LHVSPEEPRTLVT PAALRPLVLGGNQLVIIVGIVCATILLLPVLILIVKKT CGNKSSVNSTV
LVKNTKKTNP EIKEKPCHFERCEGEKHIYSPIIVREVIEEEEPSEKSEATYMTMHPVWPSLR
SDRNN SLEKKSGGGMPKTQQAF

305/330

FIGURE 305

CTATGAAGAAGCTTCCTGGAAAACAATAAGCAAAGGAAAACAAATGTGTCCCATCTCACATG
GTTCTACCCTACTAAAGACAGGAAGATCATAAACTGACAGATACTGAAATTGTAAGAGTTGG
AAACTACATTTTGCAAAGTCATTGAACTCTGAGCTCAGTTGCAGTACTCGGGAAGCC**ATG**CA
GGATGAAGATGGATACATCACCTTAAATATTAAAACCTCGGAAACCAGCTCTCGTCTCCGTTG
GCCCTGCATCCTCCTCCTGGTGGCGTGTGATGGCTTTGATTCTGCTGATCCTGTGCGTG
ATGGTTGTCTGGGCTGGTGGCTCTGGGGATTTGGTCTGTCATGCAGCGCAATTACCTACAAGA
TGAGAATGAAAATCGCACAGGAACCTGCAACAATTAGCAAAGCGCTTCTGTCAATATGTGG
TAAAACAATCAGAACTAAAGGGCACTTTCAAAGGTCATAAATGCAGCCCCCTGTGACACAAAC
TGGAGATATTATGGAGATAGCTGCTATGGGTTCTTCAGGCACAACCTAACATGGGAAGAGAG
TAAGCAGTACTGCACTGACATGAATGCTACTCTCCTGAAGATTGACAACCGGAACATTGTGG
AGTACATCAAAGCCAGGACTCATTTAATTCGTTGGGTCGGATTATCTCGCCAGAAGTCGAAT
GAGGTCTGGAAGTGGGAGGATGGCTCGGTTATCTCAGAAAATATGTTTGAGTTTTTGGAAGA
TGGAAAAGGAAATATGAATTGTGCTTATTTTCATAATGGGAAAATGCACCCTACCTTCTGTG
AGAACAAACATTATTTAATGTGTGAGAGGAAGGCTGGCATGACCAAGGTGGACCAACTACCT
TAATGCAAAGAGGTGGACAGGATAACACAGATAAGGGCTTTATTGTACAATAAAAGATATGT
ATGAATGCATCAGTAGCTGAAAAAAAAAAAAA

306/330

FIGURE 306

MQDEDGYITLNIKTRKPALVSVGPASSSWWRVMALILLILCVGMVVGLVALGIWSVMQRNYL
QDENENRTGTLQQLAKRFCQYVVKQSELKGTFKGHKCSPCDTNWRYYGDSYGFRRHNLWE
ESKQYCTDMNATLLKIDNRNIVEYIKARTHLIRWVGLSRQKSNEVWKWEDGSVISENMFEEFL
EDGKGNMNCAYFHNGKMHPTFCENKHYLMCERKAGMTKVDQLP

307/330

FIGURE 307

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCGCGATCCCGG
CCCGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCCGCGCGGGAGCCGGACCGC
CGCCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA
GCCCCGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCGGACAAGCAG
CGGAGGAGAAGGAGGAGGAGGCGAACCCAGAGAGGGGCAGCAAAAGAAGCGGTGGTGGTGGG
CGTCGTGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCG
AGCGCGAGAAATCCAACGCCTGCAAGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGC
GACAAAAACAAGTTAAATGTCTTTTCCCGGGTCAAACCTCTTCGGCTCCAAGAAGAGGCGCAG
AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC
ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT
CTGTTTAACCTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT
GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACTTTTACACCTGAGTGCA
AATTCAAAGAATCAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAG
CAGCAGTCAGGCCGAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAA
CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT
ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC
AAGAGCAGAAGTGTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC
G**TAG**CCAGTGAGGGCAAAAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAAT
TCTTCTAGCAGTCCTTCACCCAAAAGTTCAAATTTGTCAGTGACATTTACCAAACAAACAGG
CAGAGTTCACTATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

308/330

FIGURE 308

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498

><subunit 1 of 1, 245 aa, 1 stop

><MW: 27564, pI: 10.18, NX(S/T): 1

MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGKTSCDKNKLNVFSRVKLFSGSKRRRRRP
EPQLKGIVTKLYSRQGYHLQLQADGTIDGTKDEDSTYTLFNLI PVGLRVVAIQGVQTKLYLA
MNSEG YLYTSELFTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVK
KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKRSVSGVLNGGKSMHNEST

N-glycosylation site.

amino acids 242-246

Glycosaminoglycan attachment site.

amino acids 165-169, 218-222

Tyrosine kinase phosphorylation site.

amino acids 93-100

N-myristoylation site.

amino acids 87-93, 231-237

ATP/GTP-binding site motif A (P-loop).

amino acids 231-239

HBGF/FGF family proteins

amino acids 78-94, 102-153

309/330

FIGURE 309

CCAGGATGGAGCTGGGGCCTGTATAGCCATATTATTGTTCTATGCTACTAGACATGGGGGGG
ACTTGGTGAAAAAGGTATTATCCAGCCAGAGGGTCTGGGAGCCCTGTCTTACTGAACCTGGG
CAACCTGGATATTCTGAGACATATTTTGGGGGGATTTCAGTGAAAAAGTGGGGGATCCCCT
CCATTTAGAGTGTAGCAAAGGAAAAAACACCAAGGTTGGGTTCTTCCTGACATTGGCAGTG
CCCCAGTAGGGGTGGGATGAGCGAATATTCCCAAAGCTAAAGTCCACACCCCTGTAGATTAC
AAGAGTGGATTTGGCAGGAGTGTGCCCCAAAATACAGTGGAAGGTGCCTGAAGATATTTAA
ACCACGTCTTGGAATTTAGTGGGTCTTGGCTTTGGGATAGGTGAAGTGAGGACAGACACTG
GAGAGGAGGGAAAGGGGACGTTTTCAATAGGAGGCAAAACCTCGAGGGTGGGATCCACTGAGG
AGTACATAGGCTGCTGGATCTGGTGGAGCCAGCACTGGGCCCACGGGTGGTAACCTGGCTGCT
GTGGAGGGGGGTACGTGAGGGGGGGGTCTGGGGCTTATCCTCAGGTCCTGTGGGTGGGGCAG
CGAGTCGGGGCCTGAGCGTCAAGAGCATGCCCTAGTGAGCGGGCTCCTCTGGGGGAGCCCAG
CGCGCTCCGGGCGCCTGCCGGTTTGGGGGTGTCTCCTCCCGGGGCGCT**ATG**GCGGCGCTGGC
CAGTAGCCTGATCCGGCAGAAGCGGGAGGTCCGCGAGCCCGGGGGCAGCCGGCCGGTGTGCG
CGCAGCGGCGCGTGTGTCCCCGCGGCACCAAGTCCCTTTGCCAGAAGCAGCTCCTCATCCTG
CTGTCCAAGGTGCGACTGTGCGGGGGGCGGCCCCGCGCGGCCGGACCGCGGCCCGGAGCCTCA
GCTCAAAGGCATCGTCACCAAACCTGTTCTGCCGCCAGGGTTTCTACCTCCAGGCGAATCCCG
ACGGAAGCATCCAGGGCACCCAGAGGATACCAGCTCCTTCACCCACTTCAACCTGATCCCT
GTGGGCCTCCGTGTGGTCACCATCCAGAGCGCCAAGCTGGGTCACTACATGGCCATGAATGC
TGAGGGACTGCTCTACAGTTCGCCGCATTTACAGCTGAGTGTGCTTTAAGGAGTGTGTCT
TTGAGAATTACTACGTCCTGTACGCCTCTGCTCTCTACCGCCAGCGTCGTTCTGGCCGGGCC
TGGTACCTCGGCCTGGACAAGGAGGGCCAGGTCATGAAGGGAAACCGAGTTAAGAAGACCAA
GGCAGCTGCCCACTTTCTGCCCAAGCTCCTGGAGGTGGCCATGTACCAGGAGCCTTCTCTCC
ACAGTGTCCCCGAGGCCTCCCCTTCCAGTCCCCCTGCCCCC**TGA**AATGTAGTCCCTGGACTG
GAGGTTCCCTGCACTCCCAGTGAGCCAGCCACCACCACAACCTGT

310/330

FIGURE 310

MAALASSLIRQKREVPGGSRPVSAQRRVCPRGTKSLCQKQLLILLSKVRLCGGRPARPDR
GPEPQLKGIVTKLFCRQGFIYLQANPDGSIQGTPEDTSSFTHFNLIPVGLRVVTIQSAKLGHY
MAMNAEGLLYSSPHFTAECRFKECVFENYYVLYASALYRQRRSGRAWYLGLDKEGQVMKGNR
VKKTKAAAHFLPKLLEVAMYQEPSLHSVPEASPSPPAP

Tyrosine kinase phosphorylation site:

amino acids 199-207

N-myristoylation sites:

amino acids 54-60, 89-95, 131-137

HBGF/FGF family signature:

amino acids 131-155

311/330

FIGURE 311

ATGGCCGCGGCCATCGCTAGCGGCTTGATCCGCCAGAAGCGGCAGGCGCGGGAGCAGCACTG
GGACCGGCCGTCTGCCAGCAGGAGGCGGAGCAGCCCCAGCAAGAACCGCGGGCTCTGCAACG
GCAACCTGGTGGATATCTTCTCCAAAGTGCGCATCTTCGGCCTCAAGAAGCGCAGGTTGCGG
CGCCAAGATCCCCAGCTCAAGGGTATAGTGACCAGGTTATATTGCAGGCAAGGCTACTACTT
GCAAATGCACCCCGATGGAGCTCTCGATGGAACCAAGGATGACAGCACTAATTCTACACTCT
TCAACCTCATAACAGTGGGACTACGTGTTGTTGCCATCCAGGGAGTGAAAACAGGGTTGTAT
ATAGCCATGAATGGAGAAGGTTACCTCTACCCATCAGAACTTTTTACCCCTGAATGCAAGTT
TAAAGAATCTGTTTTTGAAAATTATTATGTAATCTACTCATCCATGTTGTACAGACAACAGG
AATCTGGTAGAGCCTGGTTTTTTGGGATTAAATAAGGAAGGGCAAGCTATGAAAGGGAACAGA
GTAAAGAAAACCAAACCAGCAGCTCATTTTCTACCCAAGCCATTGGAAGTTGCCATGTACCG
AGAACCATCTTTGCATGATGTTGGGGAAACGGTCCCGAAGCCTGGGGTGACGCCAAGTAAAA
GCACAAGTGCGTCTGCAATAATGAATGGAGGCAAACCAGTCAACAAGAGTAAGACAACA**TAG**

312/330

FIGURE 312

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503

><subunit 1 of 1, 247 aa, 1 stop

><MW: 27702, pI: 10.36, NX(S/T): 2

MAAAIASGLIRQKRQAREQHWDRPSASRRRSSPSKNRGLCNGNLVDIFSKVRI FGLKKRRRLR
RQDPQLKGIVTRL YCRQGYLQMHDPD GALDGTKDDSTNSTLFNLIPVGLRVVAIQGVKTGLY
IAMNGEGYLYPSELF TPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR
VKKTKPAAHFLPKPLEVAMYREPSLHDVGETVPKPGVTPSKSTSASAIMNGGKPVNKS KTT

N-glycosylation site.

amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 28-32, 29-33

Tyrosine kinase phosphorylation site.

amino acids 199-207

N-myristoylation site.

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

HBGF/FGF family proteins.

amino acids 104-155, 171-198

313/330

FIGURE 313

GGGGAGAGGAATTGACCATGTAAAAGGAGACTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAAAT
GAAGGATGCAGGACGCAGCTTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGGAAGA
ACGAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATG
AAATAAACAGAGTTAGACCCGCGGGGGTGGTGTGTTCTGACATAAATAAATAATCTTAAAGCAGCTGTTCCC
CTCCCCACCCCCAAAAAAGGATGATTGGAATGAAGAACCGAGGATTCACAAAGAAAAAAGTATGTTCAATTT
TTCTCTATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAA
AGAAGTGGTGTGGTGGTGTTCCTTTCTTTTGAATTTCCACAAAGAGGAGAGGAAATTAATAATACATCTGC
AAAGAAATTTTCAAGAGAAGAAAAGTTGACCGCGGCAGATTGAGGCATTGATTGGGGGAGAGAAACCAGCAGAGCA
CAGTTGGATTTGTGCCTATGTTGACTAAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTT
TTTTTAAATTTTTATTCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTCTTAACCACCTGGATTTCCATCT
GGATGTTGCTGTGATCAGTCTGAAATACAACCTGTTTGAATTCAGAAGGACCAACACCAGATAAATTATGAATG
TTGAACAAGATGACCTTACATCCACAGCAGATAATGATAGGTCTAGGTTTAACAGGGCCCTATTTGACCCCT
GCTTGTGGTGTGCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGC GGCTCAGACCTGCCCTTCTGTGT
GCTCCTGCAGCAACCAGTTCAGCAAGGTGATTTGTGTTGCGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCC
ACCAACACACGGCTGCTGAACCTCCATGAGAACC AAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAG
GCACTTGGAATCCTACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGA
ACCTCAACACTCTGGAACCTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAA
CTGAAGGAGCTCTGGTTGCGAAACAACCCCATTTGAAAGCATCCCTTCTTATGCTTTTAACAGAATTCCTTCTTT
GCGCCGACTAGACTTAGGGGAATTGAAAAGACTTTTATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAAC
TGAGGTATTTGAACCTTGCCATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAAACCTAGATGAG
CTGGATCTTTCTGGGAATCATTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGCACCTTCAAAAAC
GTGGATGATACAGTCCCAGATTCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCA
ACCTGGCACACAATAATCTAACATTACTGCCTCATGACCTCTTCACTCCCTTGATCATCTAGAGCGGATACAT
TTACATCACAACCTTGGAACCTGTAACCTGTGACATACTGTGGCTCAGCTGGTGGATAAAAAGACATGGCCCCCTC
GAACACAGCTTGTTGTGCCCCGGTGTAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGA
ATTACTTCACATGCTATGCTCCGGTGATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCT
GAGCTGAAATGTCGGGCCTCCACATCCCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACA
TGGGGCGTACAAAGTGCGGATAGCTGTGCTCAGTGATGGTACGTTAAATTTACAAATGTAACCTGTGCAAGATA
CAGGCATGTACACATGTATGGTGAGTAATTCCGTTGGGAATACTACTGCTTCAGCCACCTGAATGTTACTGCA
GCAACCACTACTCCTTTCTCTTACTTTTCAACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACG
GACCACAGATAACAATGTGGGTCCCACTCCAGTGCTGACTGGGAGACCACCAATGTGACCACCTCTCTCACAC
CACAGAGCACAAGGTCGACAGAGAAAACCTTCACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATT
GATGAGGTCATGAAGACTACCAAAATCATCATTTGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCT
GGTCATTTTCTACAAGATGAGGAAGCAGCACCATCGGCAAAACCATCACGCCCCAACAAAGGACTGTTGAAATTA
TTAATGTGGATGATGAGATTACGGGAGACACACCCATGGAAAGCCACCTGCCATGCCTGCTATCGAGCATGAG
CACCTAAATCACTATAACTCATACAAATCTCCCTTCAACCACACAACAACAGTTAACACAATAAATTCATACA
CAGTTCAGTGCATGAACCGTTATTGATCCGAATGAACTCTAAAGACAATGTACAAGAGACTCAAATCTAAACA
TTTACAGAGTTACAAAAACAACAATCAAAAAAAGACAGTTTATTAAAAATGACACAAATGACTGGGCTAA
ATCTACTGTTTCAAAAAAGTGCTTTTACAAAAAACA AAAAGAAAGAAATTTATTTATTA AAAATTCATTG
TGATCTAAAGCAGACAAAA

314/330

FIGURE 314

MLNKMTLHPQQIMIGPRFNRALFDPLLVLALLQLLVVAGLVRAQTCPSVCSCSNQFSKVIC
VRKNLREVPDGISTNTRLLNLHENQIQIIVKNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA
NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS
YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLELDLSDGNHLSAIRPGSFQGLMHLQKL
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNCDIL
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE
LKCRASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVNSVGN
TTASATLNVTAATTTFFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ
STRSTKTFITIPVTDINSIGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHRQN
HHAPTRTVEIINVDEITGDTPMESHLPMPAIEHEHLNHYSYKSPFNHTTTVNTINSIHSS
VHEPLLIRMNSKDNVQETQI

Signal sequence:

amino acids 1-44

Transmembrane domain:

amino acids 523-543

N-glycosylation site.amino acids 278-282, 364-368, 390-394, 412-416, 415-419,
434-438, 442-446, 488-492, 606-610**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 183-187

Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

N-myristoylation site.amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,
391-397, 422-428, 433-439, 531-537

315/330

FIGURE 315

GCGCCGGGAGCCCATCTGCCCCCAGGGGACGCGGGGCGCGGGGCGGGCTCCCGCCCCGGCACAT
GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCGCGCCAGCTCGCCCGAGGTCCGTCCGA
GGCGCCCCGGCCGCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCCGCGTCCGGGGATC
GGG**ATG**TCCCTCCTCCTTCTCCTCTTGCTAGTTTCTCTACTATGTTGGAACCTTGGGGACTCA
CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAA
GTGGTGATCACTTACTCCAGTCGTCATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG
AGTGGCCTTTGCTTCCAATTTCTGGCAGGAGATGCCTCCTTGACAGATTGAACCTCTGAAGC
CCAGTGATGAGGGCCGGTACACCTGTAAGGTAAAGAATTCAGGGCGCTACGTGTGGAGCCAT
GTCATCTTAAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC
AGAAGGAAGTGACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT
ACTGGCAGCGAATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATT
GACTACAACCACCCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCTACTCTGGACTGTA
CCAGTGACACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAAGTGTACAGT
ATGTACAAAGCATCGGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG
ATTTTCTCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA
GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT
CCTCTTCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCTCTCCACTCGCTCCACAGCAAAT
AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCCGAGCCAGGGCTGGCCAC
CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
CTAATCTGACCAAAGCAGAAACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAA
ACGGTCT**TGA**ATTACAATGGACTTGACTCCCACGCTTTCTTAGGAGTCAGGGTCTTTGGACTC
TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA
GTGAGCATTGCACGGAACAGATTCAGATGAGCATTTTCTTATAACAATACCAAACAAGCAAA
AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGG
AAAGCAGGAGTCCAAATCTATTTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG
AGGTGAATATACCTAAAACTTTTAATGTGGGATATTTTGTATCAGTGCTTTGATTCACAATT
TTCAAGAGGAAATGGGATGCTGTTTGTAATTTTCTATGCATTTCTGCAAACCTTATTGGATT
ATTAGTTATTAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
TGAGCTAACCCTTCTAAGAACTCCAAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC
TTCATTTGTCATAAGGTTTGGATATTAATTTCAAGGGGAGTTGAAATAGTGGGAGATGGAGA
AGAGTGAATGAGTTTCTCCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAAATAAC
TATGAAAGGAGACAAAAATTTGTGACAAAGGATTGTGAAGAGCTTCCATCTTCATGATGTT
ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCTCAAAT
CAGATGCCTCTAAGGACTTTCCTGCTAGATATTTCTGGAAGGAGAAAATACAACATGTCATT
TATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA
CCCAACATAACATTATAGTCTCTTCTTCTGAGAAAATGTGAAACCAGAATTGCAAGACTGG
GTGGACTAGAAAGGGAGATTAGATCAGTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA
TGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

316/330

FIGURE 316

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419

><subunit 1 of 1, 373 aa, 1 stop

><MW: 41281, pI: 8.33, NX(S/T): 3

MSLLLLLLLLVSYYVGTGLGTHTEIKRVAEEKVTLPCHHQGLPEKDTLDIEWLLTDNEGNQKV
VITYSSRHVYNNLTTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPKCELEGELTEGSDTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANS
ASRSQRTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFQTV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

FIGURE 317

[illegible]

318/330

FIGURE 318

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA82361

><subunit 1 of 1, 352 aa, 1 stop

><MW: 38938, pI: 7.86, NX(S/T): 3

MALLLCFVLLCGVVDFAFARSLSITTPPEEMIEKAKGETAYLPCKFTLSPEDQGPLDIEWLISPA
DNQKVDQVIILYSGDKIYDDYYPDLKGRVHFTSNDLKSGDASINVTNLQLSDIGTYQCKVKK
APGVANKKIHLLVVLVKPSGARCIVDGSEEIGSDFKIKCEPKEGSLPLQYEWQKLSDSQKMPT
SWLAEMTSSVISVKNASSEYSGTYSCTVRNRVGSQCLLRLNVVPPSNKAGLIAGAIIGTLL
ALALIGLIIFCCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHSSLGSMSPSNM
EGYSKTQYNQVPSSEDFERTPQSPTLPPAKFKYPYKTDGITVV

Signal sequence.

amino acids 1-19

Transmembrane domain:

amino acids 236-257

N-glycosylation sites.

amino acids 106-110, 201-205, 298-302

Tyrosine kinase phosphorylation sites.

amino acids 31-39, 78-85, 262-270

N-myristoylation sites.amino acids 116-122, 208-214, 219-225, 237-243, 241-247,
245-251, 296-302**Myelin P0 protein.**

amino acids 96-125

319/330

FIGURE 319

TGAAATGACTTCCACGGCTGGGACGGGAACCTTCCACCCACAGCTATGCCTCTGATTGGTGA
ATGGTGAAGGTGCCTGTCTAACTTTTCTGTAAAAAGAACCAGCTGCCTCCAGGCAGCCAGCC
CTCAAGCATCACTTACAGGACCAGAGGGACAAGACATGACTGTGATGAGGAGCTGCTTTTCGC
CAATTTAACACCAAGAAGAATTGAGGCTGCTTGGGAGGAAGGCCAGGAGGAACACGAGACTG
AGAG**ATGA**ATTTTCAACAGAGGCTGCAAAGCCTGTGGACTTTAGCCAGACCCTTCTGCCCTC
CTTTGCTGGCGACAGCCTCTCAAATGCAGATGGTTGTGCTCCCTTGCCTGGGTTTTACCCTG
CTTCTCTGGAGCCAGGTATCAGGGGGCCAGGGCCAAGAATTCCACTTTGGGCCCTGCCAAGT
GAAGGGGGTTGTTCCCCAGAACTGTGGGAAGCCTTCTGGGCTGTGAAAGACACTATGCAAG
CTCAGGATAACATCACGAGTGCCCGGCTGCTGCAGCAGGAGGTCTGCAGAACGTCTCGGAT
GCTGAGAGCTGTTACCTTGTCCACACCCTGCTGGAGTTCTACTTGAAAAGTGTTCAAAAA
CCACCACAATAGAACAGTTGAAGTCAGGACTCTGAAGTCATTCTCTACTCTGGCCAACAAC
TTGTTCTCATCGTGTCAAACTGCAACCCAGTCAAGAAAATGAGATGTTTTCCATCAGAGAC
AGTGCACACAGGCGGTTTTCTGCTATTCCGGAGAGCATTCAAACAGTTGGACGTAGAAGCAGC
TCTGACCAAAGCCCTTGGGGAAGTGGACATTCTTCTGACCTGGATGCAGAAATTCTACAAGC
TC**TGA**ATGTCTAGACCAGGACCTCCCTCCCCCTGGCACTGGTTTGTTCCTGTGTCAATTTCA
AACAGTCTCCCTTCCTATGCTGTTCACTGGACACTTCACGCCCTTGGCCATGGGTCCCATT
TTGGCCCAGGATTATTGTCAAAGAAGTCATTCTTTAAGCAGCGCCAGTGACAGTCAGGGGAAG
GTGCCTCTGGATGCTGTGAAGAGTCTACAGAGAAGATTCTTGTATTTATTACAACTCTATTT
AATTAATGTCAGTATTTCAACTGAAGTTCTATTTATTTGTGAGACTGTAAGTTACATGAAGG
CAGCAGAATATTGTGCCCCATGCTTCTTTACCCCTCACAATCCTTGCCACAGTGTGGGGCAG
TGGATGGGTGCTTAGTAAGTACTTAATAAACTGTGGTGCTTTTTTTGGCCTGTCTTTGGATT
GTTAAAAAACAGAGAGGGATGCTTGGATGTAAACTGAACTTCAGAGCATGAAAATCACACT
GTCTTCTGATATCTGCAGGGACAGAGCATTGGGGTGGGGGTAAGGTGCATCTGTTTGAAAAG
TAAACGATAAAATGTGGATTAAAGTGCCCAGCACAAAGCAGATCCTCAATAAACATTTTCATT
TCCCACCCACACTCGCCAGCTCACCCCATCATCCCTTTCCCTTGGTGCCCTCCTTTTTTTTT
TATCCTAGTCATTCTTCCCTAATCTTCCACTTGAGTGTCAAGCTGACCTTGCTGATGGTGAC
ATTGCACCTGGATGTACTATCCAATCTGTGATGACATTCCCTGCTAATAAAAGACAACATAA
CTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

320/330

FIGURE 320

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002

><subunit 1 of 1, 206 aa, 1 stop

><MW: 23799, pI: 9.12, NX(S/T): 3

MNFQQRLQSLWTLARPFCPPLLATASQMOMVVLPCLGFTLLLWSQVSGAQGQEFHFGPCQVK
GVVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLEVHTLLEFYLLKTVFKNH
HNRTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRFLLFRRAFKQLDVEAAL
TKALGEVDILLTWMQKFYKL

Signal sequence:

amino acids 1-42

N-glycosylation sites.

amino acids 85-89, 99-103, 126-130

321/330

FIGURE 321

AAGGAGCAGCCCGCAAGCACCAAGTGAGAGGC**ATGA**AGTTACAGTGTGTTTCCCTTTGGCTC
CTGGGTACAATACTGATATTGTGCTCAGTAGACAACCACGGTCTCAGGAGATGTCTGATTTC
CACAGACATGCACCATATAGAAGAGAGTTTCCAAGAAATCAAAAGAGCCATCCAAGCTAAGG
ACACCTTCCCAAATGTCACTATCCTGTCCACATTGGAGACTCTGCAGATCATTAAGCCCTTA
GATGTGTGCTGCGTGACCAAGAACCTCCTGGCGTTCTACGTGGACAGGGTGTTC AAGGATCA
TCAGGAGCCAAACCCCAAAATCTTGAGAAAAATCAGCAGCATTGCCAACTCTTTCCTCTACA
TGCAGAAAACCTCTGCGGCAATGTCAGGAACAGAGGCAGTGTCACTGCAGGCAGGAAGCCACC
AATGCCACCAGAGTCATCCATGACAACTATGATCAGCTGGAGGTCCACGCTGCTGCCATTAA
ATCCCTGGGAGAGCTCGACGTCTTTCTAGCCTGGATTAATAAGAATCATGAAGTAATGTTCT
CAGCT**TGA**TGACAAGGAACCTGTATAGTGATCCAGGGATGAACACCCCCTGTGCGGTTTACT
GTGGGAGACAGCCCACCTTGAAGGGGAAGGAGATGGGGAAGGCCCTTGCAGCTGAAAGTCC
CACTGGCTGGCCTCAGGCTGTCTTATTCGCTTGAAAATAGGC AAAAAGTCTACTGTGGTAT
TTGTAATAAACTCTATCTGCTGAAAGGGCCTGCAGGCCATCCTGGGAGTAAAGGGCTGCCTT
CCCATCTAATTTATTGTAAAGTCATATAGTCCATGTCTGTGATGTGAGCCAAGTGATATCCT
GTAGTACACATTGTACTGAGTGGTTTTTCTGAATAAATTCCATATTTTACCTATGA

322/330

FIGURE 322

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282
><subunit 1 of 1, 177 aa, 1 stop
><MW: 20452, pI: 8.00, NX(S/T): 2
MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST
LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQ
RQCHCRQEATNATRVIHNDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA
```

Signal sequence:

amino acids 1-18

N-glycosylation sites.

amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation site.

amino acids 24-30

Actinin-type actin-binding domain signature 1.

amino acids 159-169

323/330

FIGURE 323

CCCGTGCCAAGAGTGACGTAAGTACCGCCTATAGAGTCTATAGGCCCACTTGGCTTCGTTAG
AACGCGGCTACAATTAATACATAACCTTATGTATCATAACATACGATTTAGGTGACACTAT
AGAATAACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGGTCCAACCTGCACCTC
GGTTCTATCGATAATCTCAGCACCCAGCCACTCAGAGCAGGGCACG**ATG**TTGGGGGCCCCGCCT
CAGGCTCTGGGTCTGTGCCTTGTGCAGCGTCTGCAGCATGAGCGTCCTCAGAGCCTATCCCA
ATGCCTCCCCACTGCTCGGCTCCAGCTGGGGTGGCCTGATCCACCTGTACACAGCCACAGCC
AGGAACAGCTACCACCTGCAGATCCACAAGAATGGCCATGTGGATGGCGCACCCCATCAGAC
CATCTACAGTGCCCTGATGATCAGATCAGAGGATGCTGGCTTTGTGGTGATTACAGGTGTGA
TGAGCAGAAGATACCTCTGCATGGATTTTCAGAGGCAACATTTTTGGATCACACTATTTTCGAC
CCGGAGAACTGCAGGTTCCAACACCAGACGCTGGAAAACGGGTACGACGTCTACCACTCTCC
TCAGTATCACTTCCTGGTCAGTCTGGGCCGGGCGAAGAGAGCCTTCTGCTGCCAGGCATGAACC
CACCCCCGTACTCCCAGTTCCTGTCCCGGAGGAACGAGATCCCCCTAATTCACCTTCAACACC
CCCATACCACGGCGGCACACCCGGAGCGCCGAGGACGACTCGGAGCGGGACCCCTGAACGT
GCTGAAGCCCCGGGCCCCGGATGACCCCGGCCCCGGCCTCCTGTTACAGGAGCTCCCGAGCG
CCGAGGACAACAGCCCGATGGCCAGTGACCCATTAGGGGTGGTCAGGGGCGGTCGAGTGAAC
ACGCACGCTGGGGGAACGGGCCCCGGAAGGCTGCCGCCCTTCGCCAAGTTCATC**TAG**GGTGC
CTGG

324/330

FIGURE 324

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA142238

><subunit 1 of 1, 251 aa, 1 stop

><MW: 27954, pI: 9.22, NX(S/T): 1

MLGARLRLWVCALCSVCSMSVLRAYPNASPLLGSSWGGLIHLYTATARNSYHLQIHKNGHVD
GAPHQTIYSALMIRSEDAGFVVITGVMSRRYLCMDFRGNIFGSHYFDPENCRFQHQTLENGY
DVIHSPQYHFLVSLGRAKRAFLPGMNPPPYSQLSRNEIPLIHFNTPIPRRHTRSAEDDSE
RDPLNLVKPRARMTPAPASCSQELPSAEDNSPMASDPLGVVRGGRVNTHAGGTGPEGCRPFA
KFI

Important features of the protein:**Signal peptide:**

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 175-179

N-myristoylation site.

amino acids 33-39, 100-106, 225-231, 229-235

HBGF/FGF family proteins

amino acids 73-124

325/330

FIGURE 325

GGAAAAGGTACCCGCGAGAGACAGCCAGCAGTTCTGTGGAGCAGCGGTGGCCGGCTAGG**ATG**
GGCTGTCTCTGGGGTCTGGCTCTGCCCCTTTTCTTCTTCTGCTGGGAGGTTGGGGTCTCTGG
GAGCTCTGCAGGCCCCAGCACCCGCGAGAGCAGACACTGCGATGACAACGGACGACACAGAAG
TGCCCGCTATGACTCTAGCACCGGGGCCACGCCGCTCTGGAAACTCAAACGCTGAGCGCTGAG
ACCTCTTCTAGGGCCTCAACCCCGAGCCGGCCCCATTCCAGAAGCAGAGACCAGGGGAGCCAA
GAGAATTTCCCCTGCAAGAGAGACCAGGAGTTTCACAAAAACATCTCCCAACTTCATGGTGC
TGATCGCCACCTCCGTGGAGACATCAGCCGCCAGTGGCAGCCCCGAGGGAGCTGGAATGACC
ACAGTTCAGACCATCACAGGCAGTGATCCCGAGGAAGCCATCTTTGACACCCTTTGCACCGA
TGACAGCTCTGAAGAGGCAAAGACACTCACAATGGACATATTGACATTGGCTCACACCTCCA
CAGAAGCTAAGGGCCTGTCCTCAGAGAGCAGTGCCTCTTCCGACGGCCCCCATCCAGTCATC
ACCCCGTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCC
GTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCCGTCAT
GGTCCCCGGGATCTGATGTCACCTCTCCTCGCTGAAGCCCTGGTGACTGTCACAAACATCGAG
GTTATTAATTGCAGCATCACAGAAATAGAAACAACAACCTTCCAGCATCCCTGGGGCCTCAGA
CATAGATCTCATCCCCACGGAAGGGGTGAAGGCCTCGTCCACCTCCGATCCACCAGCTCTGC
CTGACTCCACTGAAGCAAAACCACACATCACTGAGGTCACAGCCTCTGCCGAGACCCTGTCC
ACAGCCGGCACACAGAGTCAGCTGCACCTCATGCCACGGTTGGGACCCCACTCCCCACTAA
CAGCGCCACAGAAAGAGAAGTGACAGCACCCGGGGCCACGACCCTCAGTGGAGCTCTGGTCA
CAGTTAGCAGGAATCCCCTGGAAGAAACCTCAGCCCTCTCTGTTGAGACACCAAGTTACGTC
AAAGTCTCAGGAGCAGCTCCGGTCTCCATAGAGGCTGGGTCAGCAGTGGGCAAAACAACCTTC
CTTTGCTGGGAGCTCTGCTTCCTCCTACAGCCCCCTCGGAAGCCGCCCTCAAGAACTTCACCC
CTTCAGAGACACCGACCATGGACATCGCAACCAAGGGGCCCTTCCCCACCAGCAGGGACCCT
CTTCCTTCTGTCCCTCCGACTACAACCAACAGCAGCCGAGGGACGAACAGCACCTTAGCCAA
GATCACAACCTCAGCGAAGACCACGATGAAGCCCCAACAGCCACGCCCACGACTGCCCGGAC
GAGGCCGACCACAGACG**TGA**GTGCAGGTGAAAATGGAGGTTTCTCCTCCTGCGGCTGAGTG
TGGCTTCCCCGGAAGACCTCACTGACCCCAGAGTGGCAGAAAGGCTGATGCAGCAGCTCCAC
CGGGAACCTCCACGCCCACGCGCCTCACTTCCAGGTCTCCTTACTGCGTGTCAGGAGAGGCTA
ACGGACATCAGCTGCAGCCAGGCATGTCCCGTATGCCAAAAGAGGGTGCTGCCCCTAGCCTG
GGCCCCCACCAGACAGACTGCAGCTGCGTTACTGTGCTGAGAGGTACCCAGAAGGTTCCCATG
AAGGGCAGCATGTCCAAGCCCCCTAACCCCGAGATGTGGCAACAGGACCCTCGCTCACATCCAC
CGGAGTGTATGTATGGGGAGGGGCTTACCTGTTCCCAGAGGTGTCCTTGGACTCACCTTGG
CACATGTTCTGTGTTTCAGTAAAGAGAGACCTGATCACCCATCTGTGTGCTTCCATCCTGCA
TTAAAATTCACTCAGTGTGGCCCCAAAAAAA

326/330

FIGURE 326

MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTL
ETSSRASTPAGPIPEAETRGAKRISPARETRSFTKTSPNFMVLIATSVETSAASGSPEGAGM
TTVQTITGSDPEEAI FDTLCTDDSS EAKTLTMDILT LAHTSTEAKGLSSESSASSDGPHPV
ITPSRASESSASSDGPHPVITPSRASESSASSDGPHPVITPSWSPGSDVTLLAEALVTVTNI
EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL
STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTL SGALVTVSRNPLEETSALSVETPSY
VKVSGAAPVSIEAGSAVGKTTSFAGSSASSYS PSEAALKNFTPSETPTMDIATKGPFPTSRD
PLPSVPPTTTNSSRGTNSTLAKITTS AKTTMKPQQPRPRLPGRGRPQT

N-glycosylation sites:

amino acids 252-256, 445-449, 451-455

cAMP-and cGMP-dependent protein kinase phosphorylation site.

amino acids 84-90

Casein kinase II phosphorylation sites.

amino acids 37-41, 108-112, 131-135, 133-137, 148-152, 165-169,
246-250, 254-258, 256-260, 269-273, 283-287, 333-337, 335-339,
404-408, 414-418, 431-435

N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284, 314-
320, 349-355, 386-392, 397-403, 449-455

ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

327/330

FIGURE 327

GCGGAGCATCCGCTGCGGTCTCGCCGAGACCCCCGCGCGGATTGCGCCGGTCCTTCCCGCGG
GCGCGACAGAGCTGTCCTCGCACCTGGATGGCAGCAGGGGCGCCGGGGTCTCTCGACGCCA
GAGAGAAATCTCATCATCTGTGCAGCCTTCTTAAAGCAAACCTAAGACCAGAGGGAGGATTAT
CCTTGACCTTTGAAGACCAAACTAACTGAAATTTAAA**ATG**TTCTTCGGGGGAGAAGGGAG
CTTGACTTACACTTTGGTAATAATTTGCTTCCTGACACTAAGGCTGTCTGCTAGTCAGAATT
GCCTCAAAAAGAGTCTAGAAGATGTTGTTCATTGACATCCAGTCATCTCTTTCTAAGGGAATC
AGAGGCAATGAGCCCGTATATACTTCAACTCAAGAAGACTGCATTAATTCTTGCTGTTCAAC
AAAAACATATCAGGGGACAAAGCATGTAACTTGATGATCTTCGACACTCGAAAAACAGCTA
GACAACCCAACTGCTACCTATTTTTCTGTCCCAACGAGGAAGCCTGTCCATTGAAACCAGCA
AAAGGACTTATGAGTTACAGGATAATTACAGATTTTCCATCTTTGACCAGAAATTTGCCAAG
CCAAGAGTTACCCAGGAAGATTCTCTCTTACATGGCCAATTTTCACAAGCAGTCACTCCCC
TAGCCCATCATCACACAGATTATTCAAAGCCCACCGATATCTCATGGAGAGACACACTTTCT
CAGAAGTTTGGATCCTCAGATCACCTGGAGAACTATTTAAGATGGATGAAGCAAGTGCCCA
GCTCCTTGCTTATAAGGAAAAAGGCCATTCTCAGAGTTCACAATTTTCCTCTGATCAAGAAA
TAGCTCATCTGCTGCCTGAAAATGTGAGTGCCTCCAGCTACGGTGGCAGTTGCTTCTCCA
CATAACACCTCGGCTACTCCAAAGCCCGCCACCCTTCTACCCACCAATGCTTCAGTGACACC
TTCTGGGACTTCCCAGCCACAGCTGGCCACCACAGCTCCACCTGTAACCACTGTCACCTTCTC
AGCCTCCCACGACCCTCATTTCTACAGTTTTTACACGGGCTGCGGCTACACTCCAAGCAATG
GCTACAACAGCAGTTCTGACTACCACCTTTCAGGCACCTACGGACTCGAAAGGCAGCTTAGA
AACCATAACGTTTACAGAAATCTCCAACCTTAACCTTGAACACAGGGAATGTGTATAACCTA
CTGCACTTTCTATGTCAAATGTGGAGTCTTCCACTATGAATAAACTGCTTCCTGGGAAGGT
AGGGAGGCCAGTCCAGGCAGTTCTCCAGGGCAGTGTTCCAGAAAATCAGTACGGCCTTCC
ATTTGAAAAATGGCTTCTTATCGGGTCCCTGCTCTTTGGTGTCTGTTCTGCTGGTGATAGGCC
TCGTCCTCCTGGGTAGAATCCTTTCCGAATCACTCCGCAGGAAACGTTACTCAAGACTGGAT
TATTTGATCAATGGGATCTATGTGGACATCT**TAA**GGATGGAACCTCGGTGTCTCTTAATTCATT
TAGTAACCAGAAGCCCAAATGCAATGAGTTTCTGCTGACTTGCTAGTCTTAGCAGGAGGTTG
TATTTTGAAGACAGGAAAATGCCCCCTTCTGCTTTCCTTTTTTTTTTTGGAGACAGAGTCTT
GCTCTGTTGCCAGGCTGGAGTGCAGTAGCACGATCTCGGCTCTCACCGCAACCTCCGTCTC
CTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCTAAGTATCTGGGATTACAGGCATGTGCCA
CCACACCTGGGTGATTTTTGTATTTTGTAGTAGAGACGGGGTTTACCATGTTGGTCAGGCTG
GTCTCAAACCTCCTGACCTAGTGATCCACCCTCCTCGGCCTCCCAAAGTGCTGGGATTACAGG
CATGAGCCACCACAGCTGGCCCCCTTCTGTTTTATGTTTTGGTTTTTTGAGAAGGAATGAAGTG
GGAACCAAATTAGGTAATTTTGGGTAATCTGTCTCTAAAATATTAGCTAAAAACAAAGCTCT
ATGTAAAGTAATAAAGTATAATTGCCATATAAATTTCAAATTCAACTGGCTTTTATGCAAA
GAAACAGGTTAGGACATCTAGGTTCCAATTCATTCACATTCTTGGTTCCAGATAAAATCAAC
TGTTTTATATCAATTTCTAATGGATTTGCTTTTCTTTTTATATGGATTCCTTTAAACCTTATT
CCAGATGTAGTTCCTTCCAATTAAATATTTGAATAAATCTTTTGTTACTCAA

328/330

FIGURE 328

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45410

><subunit 1 of 1, 431 aa, 1 stop

><MW: 46810, pI: 6.45, NX(S/T): 6

MFFGGEGSLTYTLVIICFLTLRLSASQNCLKKSLEDVVIDIQSSLKGIKRGNEPVYTSTQED
CINSCCSTKNISGDKACNLMIFDTRKTARQPNCYLFFCPNEEACPLKPAKGLMSYRIITDFP
SLTRNLPSQELPQEDSLLHGQFSQAVTPLAHHHTDYSKPTDISWRDTLSQKFGSSDHLEKLF
KMDEASAQLLAYKEKGHSQSSQFSSDQEIAHLLPENVSALPATVAVASPHTTSATPKPATLL
PTNASVTPSGTSQPQLATTAPPVTTVTSPPTTLISTVFTRAAATLQAMATTAVLTTTTFQAP
TDSKGSLETIPFTEISNLTNTGNVYNPTALSMSNVESSTMNKTASWEGREASPGSSSQGSV
PENQYGLPFEKWLLIGSLLFGVLFLVIGLVLLGRILSESLRRKRYSRDYLINGIYVDI

Signal sequence.

amino acids 1-25

Transmembrane domain.

amino acids 384-405

N-glycosylation sites.

amino acids 72-76, 222-226, 251-255, 327-331, 352-356

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 415-419

Tyrosine kinase phosphorylation site.

amino acids 50-57

N-myristoylation sites.

amino acids 4-10, 48-54, 315-321

329/330

FIGURE 329

CTCCACGGTGTCCAGCGCCCAGAA**ATG**CGGCTTCTGGTCCTGCTATGGGGTTGCCTGCTGCT
CCCAGGTTATGAAGCCCTGGAGGGGCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGT
CCCTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT
GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGA
ACCTCACCCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAG
TCTTTACTGATCTCTCTGTTTCGTCTTTCCAGGACCCTGCTGTCTCTCCCTCCCCTTCTCCCAC
CTTCCAGCCTCTGGCTACAACACGCCTGCAGCCCCAAGGCAAAAGCTCAGCAAACCCAGCCCC
CAGGATTGACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
GCTGAGGCCCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
AGGAACCTCTCCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCCATGCAGC
TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCTGGTGCTGCTGAGCCTTCTGTGAGC
CGCAGGCCTGATCGCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCA
CGGAGACACAGAGGAACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCC
CCTTCCCAGGCCCCCTGAGGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGA
GCTGGGCTTCTCGAAGTTTGTCTCAGCG**TAG**GGCAGGAGGCCCTCCTGGCCAGGCCAGCAGT
GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCCCGAAAGCTTTCCACCTCAGCCTCAGAG
TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCAGGCTCTCCTCTTGATGTTCCA
GCCTGACCTAGAAGCGTTTGTGAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG
GAGACTGGGACATCCCTGATAGGTTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCA
GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACTCCTGGGC
CTCATGCCCAGTGTGCGACCCCTGCCTTCTCCCACTCCAGACCCACCTTGTCTTCCCTCCC
TGGCGTCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGATGAAGAGGAGCATGCT
GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT
GTGAAAAACGTGATTCTTGGCCCCACCAAGACCCACCAAAAACCATCTCTGGGCTTGGTGCAG
GACTCTGAATTCTAACAATGCCCAGTGAAGTGTGCACTTGAGTTTGAGGGCCAGTGGGCCTG
ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC
CAATAGATCTGCTCTGTCTGCGACACCAGATCCACGTGGGGACTCCCCTGAGGCCTGCTAAG
TCCAGGCCTTGGTCAGGTCAGGTGCACATTGCAGGATAAGCCCAGGACCGGCACAGAAGTGG
TTGCCCTTNNCCATTTGCCCTCCCTGGNCCATGCCTTCTTGCCCTTTGGAAAAAATGATGAAGA
AAACCTTGGCTCCTTCCCTTGTCTGGAAAGGGTTACTTGCCTATGGGTTCTGGTGGCTAGAGA
GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG
ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGCGGGGTGGTGGTAAAGTA
GCACAATACTATTTTTTTTCTTTTTCCATTATTATTGTTTTTTAAGACAGAATCTCGTGCT
GCTGCCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACCTCCGCCTCCTGGGTTCAAGTGATT
CTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACCACACCTGGCTAATT
TTTGTACTTTTAGTAGAGATGGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACCTCCTGAC
CTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCGGGATTACAGGCATGAGCCACTGTG
TCTGGCCCTATTTCTTTAAAAAGTGAAATTAAGAGTTGTTTCAGTATGCAAACTTGGAAG
ATGGAGGAGAAAAAGAAAAGGAAGAAAAAAATGTCACCCATAGTCTCACCAGAGACTATCAT
TATTTTCGTTTTGTGTACTTCCCTTCCACTCTTTTCTTCTTACATAATTTGCCGGTGTCTT
TTTACAGAGCAATTATCTTGTATATACAACTTTGTATCCTGCCTTTTCCACCTTATCGTTCC
ATCACTTTATTCCAGCACTTCTCTGTGTTTTACAGACCTTTTTATAAATAAAATGTTTCATCA
GCTGCATAAAAAAAAAAAAAA

330/330

FIGURE 330

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALEGP EEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAEEEGQETMKGRVSIRDSRQELSLIVTLWNLTLDAGEYWCGVEKRGPD ELLISLFV
FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGA EAPPLPG
TSQYGHERTSQYTGTSPHPATSP PAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLRKEAQQTETQRNEKFWLSRLTAE EKEAPSQAPEGD
VISM PPLHTSEEELGFSKFVSA

Important features:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 00/08439

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/47 C07K14/705 C12N15/62 C07K16/18
 G01N33/53 A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K G01N A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

STRAND

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 39448 A (HUMAN GENOME SCIENCES INC) 11 September 1998 (1998-09-11) see passages relating to gene 174 (clone H2MBF44) and the claims. ---	1-28
X	WO 98 42741 A (GENETICS INST) 1 October 1998 (1998-10-01) see passages relating to clone ck181_7 and the claims. ---	1-28
A	EP 0 834 563 A (SMITHKLINE BEECHAM CORP) 8 April 1998 (1998-04-08) the whole document ---	
A	WO 97 07198 A (GENETICS INST) 27 February 1997 (1997-02-27) the whole document ---	
	--- -/--	



Further documents are listed in the continuation of box C.



Patent family members are listed in annex.

° Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance
 "E" earlier document but published on or after the international filing date
 "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
 "O" document referring to an oral disclosure, use, exhibition or other means
 "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

16 August 2000

Date of mailing of the international search report

13.11.00

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

Smalt, R

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	YOKOYAMA-KOBAYASHI M ET AL: "A signal sequence detection system using secreted protease activity as an indicator" GENE,NL,ELSEVIER BIOMEDICAL PRESS. AMSTERDAM, vol. 163, no. 2, 3 October 1995 (1995-10-03), pages 193-196, XP004041983 ISSN: 0378-1119 the whole document	
A	--- KLEIN R D ET AL: "Selection for genes encoding secreted proteins and receptors" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA,US,NATIONAL ACADEMY OF SCIENCE. WASHINGTON, no. 93, 1 July 1996 (1996-07-01), pages 7108-7113, XP002077277 ISSN: 0027-8424 the whole document	
P,X, L	--- WO 99 63088 A (BAKER KEVIN ;CHEN JIAN (US); GENENTECH INC (US); YUAN JEAN (US); G) 9 December 1999 (1999-12-09) see passages relating to PR0281 and the claims. L: priority.	1-28
P,X	--- WO 99 46375 A (SCHMITT ARMIN ;SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BERND) 16 September 1999 (1999-09-16) see whole document, particularly passages relating to seq.ID'2 219 and 251	1-13, 17-28
P,X	--- WO 99 53040 A (SCHMITT ARMIN ;SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN BERND) 21 October 1999 (1999-10-21) page 1 -page 8 see claims. page 128	1-13, 17-22, 24-28
P,X	--- WO 00 08145 A (NOVARTIS ERFINDUNGEN VERWALTUN ;NOVARTIS AG (CH); HOLNESS CLAIRE L) 17 February 2000 (2000-02-17) the whole document	1-13, 21-28
P,X	--- WO 99 63083 A (TSURITANI KATSUKI ;ARASE SEIJI (JP); IKEDA AKIKO (JP); TAISHO PHAR) 9 December 1999 (1999-12-09) abstract see sequences.	1-13,21, 22,25-28

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 00/08439

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Claims 1-28, All partially.

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: Invention 1: claims 1-28, all partially

Isolated nucleic acid having at least 80% sequence homology to a polynucleotide sequence encoding PR0281 (seq.ID.2), vector comprising said nucleic acid, host cell comprising said vector, method of producing said protein using said host, isolated protein having at least 80% homology to PR0281, chimeric protein comprising a portion corresponding to PR0281, antibody against said protein, the isolated extracellular domain of said protein or a protein with at least 80% homology thereto, and said isolated protein lacking its signal peptide or a protein with at least 80% homology thereto.

2. Claims: Inventions 2-132: claims 1-28,
all partially and claims 4 and 13 as far as
applicable

Subject matter as defined for invention 1 above, but limited to the respective amino acid sequences and corresponding nucleic acid sequences referred to as:

2. PR0276 (seq.ID's 5/6),
3. PR0189 (Seq.ID's 7/8),
4. PR0190 (Seq.ID's 13/14),
5. PR0341 (Seq.ID's 19/20),
6. PR0180 (Seq.ID's 22/23),
7. PR0190 (Seq.ID's 13/14),
8. PR0194 (Seq.ID's 27/28),
9. PR0203 (Seq.ID's 29/30),
10. PR0290 (Seq.ID's 32/33),
11. PR0874 (Seq.ID's 35/36),
12. PR0710 (Seq.ID's 40/41),
13. PR01151 (Seq.ID's 46/47),
14. PR01281 (Seq.ID's 51/52),
15. PR0358 (Seq.ID's 56/57),
16. PR01310 (Seq.ID's 61/62),
17. PR0698 (Seq.ID's 66/67),
18. PR0732 (Seq.ID's 72/73),
19. PR01120 (Seq.ID's 83/84),
20. PR0537 (Seq.ID's 94/95),
21. PR0536 (Seq.ID's 96/97),
22. PR0535 (Seq.ID's 98/99),
23. PR0718 (Seq.ID's 102/103),
24. PR0872 (Seq.ID's 112/113),
25. PR01063 (Seq.ID's 114/115),
26. PR0619 (Seq.ID's 116/117),
27. PR01188 (Seq.ID's 123/124),
28. PR0784 (Seq.ID's 134/135),
29. PR0783 (Seq.ID's 137/138),
30. PR0820 (Seq.ID's 145/146),

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

31. PR01080 (Seq.ID's 147/148),
32. PR01079 (Seq.ID's 150/151),
33. PR0793 (Seq.ID's 152/153),
34. PR01016 (Seq.ID's 155/156),
35. PR01013 (Seq.ID's 157/158),
36. PR0937 (Seq.ID's 159/160),
37. PR0842 (Seq.ID's 164/165),
38. PR0839 (Seq.ID's 166/167),
39. PR01180 (Seq.ID's 168/169),
40. PR01134 (Seq.ID's 170/171),
41. PR0830 (Seq.ID's 174/175),
42. PR01115 (Seq.ID's 176/177),
43. PR01277 (Seq.ID's 178/179),
44. PR01135 (Seq.ID's 180/181),
45. PR0828 (Seq.ID's 188/189),
46. PR01009 (Seq.ID's 193/194),
47. PR01007 (Seq.ID's 196/197),
48. PR01056 (Seq.ID's 198/199),
49. PR0826 (Seq.ID's 200/201),
50. PR0819 (Seq.ID's 202/203),
51. PR01006 (Seq.ID's 204/205),
52. PR01112 (Seq.ID's 206/207),
53. PR01074 (Seq.ID's 208/209),
54. PR01005 (Seq.ID's 210/211),
55. PR01073 (Seq.ID's 212/213),
56. PR01152 (Seq.ID's 215/216),
57. PR01136 (Seq.ID's 218/219),
58. PR0813 (Seq.ID's 220/221),
59. PR0809 (Seq.ID's 222/223),
60. PR0791 (Seq.ID's 224/225),
61. PR01004 (Seq.ID's 226/227),
62. PR01111 (Seq.ID's 228/229),
63. PR01344 (Seq.ID's 230/231),
64. PR01109 (Seq.ID's 235/236),
65. PR01383 (Seq.ID's 240/241),
66. PR01003 (Seq.ID's 245/246),
67. PR01108 (Seq.ID's 247/248),
68. PR01137 (Seq.ID's 249/250),
69. PR01138 (Seq.ID's 252/253),
70. PR01054 (Seq.ID's 255/256),
71. PR0994 (Seq.ID's 257/258),

3. Claim : 1

72. PR0812 (Seq.ID's 259/260),
73. PR01069 (Seq.ID's 261/262),
74. PR01129 (Seq.ID's 263/264),
75. PR01068 (Seq.ID's 265/266),
76. PR01066 (Seq.ID's 267/268),
77. PR01184 (Seq.ID's 269/270),
78. PR01360 (Seq.ID's 271/272),
79. PR01029 (Seq.ID's 273/274),
80. PR01139 (Seq.ID's 275/276),
81. PR01309 (Seq.ID's 277/278),

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

82. PR01028 (Seq.ID's 280/281),
83. PR01027 (Seq.ID's 282/283),
84. PR01107 (Seq.ID's 284/285),
85. PR01140 (Seq.ID's 286/287),
86. PR01106 (Seq.ID's 288/289),
87. PR01291 (Seq.ID's 290/291),
88. PR01105 (Seq.ID's 292/293),
89. PR0511 (Seq.ID's 294/295),
90. PR01104 (Seq.ID's 296/297),
91. PR01100 (Seq.ID's 298/299),
92. PR0836 (Seq.ID's 300/301),
93. PR01141 (Seq.ID's 302/303),
94. PR01132 (Seq.ID's 308/309),
95. PR01346 (Seq.ID's 313/314),
96. PR01131 (Seq.ID's 318/319),
97. PR01281 (Seq.ID's 325/326),
98. PR01064 (Seq.ID's 333/334),
99. PR01379 (Seq.ID's 339/340),
100. PR0844 (Seq.ID's 344/345),
101. PR0848 (Seq.ID's 346/347),
102. PR01097 (Seq.ID's 348/349),
103. PR01153 (Seq.ID's 350/351),
104. PR01154 (Seq.ID's 352/353),
105. PR01182 (Seq.ID's 356/357),
106. PR01155 (Seq.ID's 358/359),
107. PR01156 (Seq.ID's 360/361),
108. PR01098 (Seq.ID's 362/363),
109. PR01127 (Seq.ID's 364/365),
110. PR01126 (Seq.ID's 366/367),
111. PR01125 (Seq.ID's 368/369),
112. PR01186 (Seq.ID's 370/371),
113. PR01198 (Seq.ID's 372/373),
114. PR01158 (Seq.ID's 374/375),
115. PR01159 (Seq.ID's 376/377),
116. PR01124 (Seq.ID's 378/379),
117. PR01287 (Seq.ID's 380/381),
118. PR01312 (Seq.ID's 386/387),
119. PR01192 (Seq.ID's 388/389),
120. PR01160 (Seq.ID's 393/394),
121. PR01187 (Seq.ID's 398/399),
122. PR01185 (Seq.ID's 400/401),
123. PR01345 (Seq.ID's 402/403),
124. PR01245 (Seq.ID's 407/408),
125. PR01358 (Seq.ID's 409/410),
126. PR01195 (Seq.ID's 411/412),
127. PR01270 (Seq.ID's 413/414),
128. PR01271 (Seq.ID's 415/416),
129. PR01375 (Seq.ID's 417/418),
130. PR01385 (Seq.ID's 419/420),
131. PR01384 (Seq.ID's 423/424),
132. PR09828 (Seq.ID's 510/511).

For the sake of conciseness, the first subject matter is explicitly defined, the other subject matters are defined by analogy thereto.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

4. Claims: Invention 133: claims 1-36,69-74,99-102,
all partially and claims 4 and 13 as far as
applicable

Isolated nucleic acid having at least 80% sequence homology to a polynucleotide sequence encoding PR0943 (seq.ID.118), vector comprising said nucleic acid, host cell comprising said vector, method of producing said protein using said host, isolated protein having at least 80% homology to PR0943, chimeric protein comprising a portion corresponding to PR0943, antibody against said protein, the isolated extracellular domain of said protein or a protein with at least 80% homology thereto, and said isolated protein lacking its signal peptide or a protein with at least 80% homology thereto. Method for detecting PR0183, PR0184 or PR0185 through interaction with PR0943 and vice versa, method of linking a bioactive molecule to a cell expressing PR0943 using PR0183, PR0184 or PR0185 as binding ligands and vice versa, and method of modulating the activity of PR0943 by binding with PR0183, PR0184 or PR0185, and vice versa.

5. Claims: Inventions 134-136: claims 1-36,69-74,99-102,
all partially and claims 4 and 13 as far as
applicable

Isolated nucleic acid having at least 80% sequence homology to a polynucleotide sequence encoding PR0183 (seq.ID.495), PR0184 (seq.ID.497) or PR0185 (seq.ID.499), vector comprising said nucleic acid, host cell comprising said vector, method of producing said protein using said host, isolated protein having at least 80% homology to said protein, chimeric protein comprising a portion corresponding to said protein, antibody against said protein, the isolated extracellular domain of said protein or a protein with at least 80% homology thereto, and said isolated protein lacking its signal peptide or a protein with at least 80% homology thereto. Method for detecting PR0183, PR0184 or PR0185 through interaction with PR0943 and vice versa, method of linking a bioactive molecule to a cell expressing PR0943 using PR0183, PR0184 or PR0185 as binding ligands and vice versa, and method of modulating the activity of PR0943 by binding with PR0183, PR0184 or PR0185, and vice versa, whereby invention 134 is limited to PR0183, invention 135 is limited to PR0184, and invention 136 is limited to PR0185.

6. Claims: Invention 137: claims 1-28,37-44,75-80,103-106,
all partially and claims 4 and 13 as far as
applicable

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Isolated nucleic acid having at least 80% sequence homology to a polynucleotide sequence encoding PR0331 (seq.ID.501), vector comprising said nucleic acid, host cell comprising said vector, method of producing said protein using said host, isolated protein having at least 80% homology to PR0331, chimeric protein comprising a portion corresponding to PR0331, antibody against said protein, the isolated extracellular domain of said protein or a protein with at least 80% homology thereto, and said isolated protein lacking its signal peptide or a protein with at least 80% homology thereto. Method for detecting PR01133 through interaction with PR0331 and vice versa, method of linking a bioactive molecule to a cell expressing PR0331 using PR01133 as binding ligands and vice versa, and method of modulating the activity of PR0331 by binding with PR01133, and vice versa.

7. Claims: Invention 138: claims 1-28,37-44,75-80,103-106,
all partially and claims 4 and 13 as far as
applicable

Isolated nucleic acid having at least 80% sequence homology to a polynucleotide sequence encoding PR01133 (seq.ID.129), vector comprising said nucleic acid, host cell comprising said vector, method of producing said protein using said host, isolated protein having at least 80% homology to said protein, chimeric protein comprising a portion corresponding to said protein, antibody against said protein, the isolated extracellular domain of said protein or a protein with at least 80% homology thereto, and said isolated protein lacking its signal peptide or a protein with at least 80% homology thereto. Method for detecting PR01133 through interaction with PR0331 and vice versa, method of linking a bioactive molecule to a cell expressing PR0331 using PR01133 as binding ligands and vice versa, and method of modulating the activity of PR0331 by binding with PR01133, and vice versa.

8. Claims: Invention 139: claims 1-28,45-52,81-86,107-110,
all partially and claims 4 and 13 as far as
applicable

Isolated nucleic acid having at least 80% sequence homology to a polynucleotide sequence encoding PR01387 (seq.ID.422), vector comprising said nucleic acid, host cell comprising said vector, method of producing said protein using said host, isolated protein having at least 80% homology to PR01387, chimeric protein comprising a portion corresponding to PR01387, antibody against said protein, the isolated extracellular domain of said protein or a protein with at

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

least 80% homology thereto, and said isolated protein lacking its signal peptide or a protein with at least 80% homology thereto. Method for detecting PR0363 or PR05723 through interaction with PR01387 and vice versa, method of linking a bioactive molecule to a cell expressing PR01387 using PR0363 or PR05723 as binding ligands and vice versa, and method of modulating the activity of PR01387 by binding with PR0363 or PR05723, and vice versa.

9. Claims: Invention 140 and 141: claims 1-28,45-52,81-86, 107-110,
all partially and claims 4 and 13 as far as applicable

Isolated nucleic acid having at least 80% sequence homology to a polynucleotide sequence encoding PR0363 (seq.ID.503) or PR05723 (seq.ID.505), vector comprising said nucleic acid, host cell comprising said vector, method of producing said protein using said host, isolated protein having at least 80% homology to said protein, chimeric protein comprising a portion corresponding to said protein, antibody against said protein, the isolated extracellular domain of said protein or a protein with at least 80% homology thereto, and said isolated protein lacking its signal peptide or a protein with at least 80% homology thereto. Method for detecting PR0363 or PR05723 through interaction with PR01387 and vice versa, method of linking a bioactive molecule to a cell expressing PR01387 using PR0363 or PR05723 as binding ligands and vice versa, and method of modulating the activity of PR01387 by binding with PR0363 or PR05723, and vice versa, whereby invention 140 is limited to PR0363 and invention 141 is limited to PR05723.

10. Claims: Invention 142: claims 1-28, 53-60,87-92,111-114,
all partially and claims 4 and 13 as far as applicable

Isolated nucleic acid having at least 80% sequence homology to a polynucleotide sequence encoding PR01114 (seq.ID.183), vector comprising said nucleic acid, host cell comprising said vector, method of producing said protein using said host, isolated protein having at least 80% homology to PR01114, chimeric protein comprising a portion corresponding to PR01114, antibody against said protein, the isolated extracellular domain of said protein or a protein with at least 80% homology thereto, and said isolated protein lacking its signal peptide or a protein with at least 80% homology thereto. Method for detecting PR03301 or PR09940 through interaction with PR01114 and vice versa, method of linking a bioactive molecule to a cell expressing PR01114 using PR03301 or PR09940 as binding ligands and vice versa,

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

and method of modulating the activity of PR01114 by binding with PR03301 or PR09940, and vice versa.

11. Claims: Invention 143 and 144: claims 1-28, 53-60,87-92, 111-114,
all partially and claims 4 and 13 as far as applicable

Isolated nucleic acid having at least 80% sequence homology to a polynucleotide sequence encoding PR03301 (seq.ID.507) or PR09940 (seq.ID.509), vector comprising said nucleic acid, host cell comprising said vector, method of producing said protein using said host, isolated protein having at least 80% homology to said protein, chimeric protein comprising a portion corresponding to said protein, antibody against said protein, the isolated extracellular domain of said protein or a protein with at least 80% homology thereto, and said isolated protein lacking its signal peptide or a protein with at least 80% homology thereto. Method for detecting PR03301 or PR09940 through interaction with PR01114 and vice versa, method of linking a bioactive molecule to a cell expressing PR01114 using PR03301 or PR09940 as binding ligands and vice versa, and method of modulating the activity of PR01114 by binding with PR03301 or PR09940, and vice versa, whereby invention 143 is limited to PR03301 and invention 144 is limited to PR09940.

12. Claims: Invention 145: claims 1-28,61-69,93-98,115-118,
all partially and claims 4 and 13 as far as applicable

Isolated nucleic acid having at least 80% sequence homology to a polynucleotide sequence encoding PR01181 (seq.ID.355), vector comprising said nucleic acid, host cell comprising said vector, method of producing said protein using said host, isolated protein having at least 80% homology to PR01181, chimeric protein comprising a portion corresponding to PR01181, antibody against said protein, the isolated extracellular domain of said protein or a protein with at least 80% homology thereto, and said isolated protein lacking its signal peptide or a protein with at least 80% homology thereto. Method for detecting PR07170, PR0361 or PR0846 through interaction with PR01181 and vice versa, method of linking a bioactive molecule to a cell expressing PR01181 using PR07170, PR0361 or PR0846 as binding ligands and vice versa, and method of modulating the activity of PR01181 by binding with PR07170, PR0361 or PR0846, and vice versa.

13. Claims: Inventions 146-148: claims 1-28,61-69,93-98,

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

115-118,
all partially and claims 4 and 13 as far as
applicable

Isolated nucleic acid having at least 80% sequence homology to a polynucleotide sequence encoding PR07170 (seq.ID.513), PR0361 (seq.ID.515) or PR0846 (seq.ID.517), vector comprising said nucleic acid, host cell comprising said vector, method of producing said protein using said host, isolated protein having at least 80% homology to said protien, chimeric protein comprising a portion corresponding to said protein, antibody against said protein, the isolated extracellular domain of said protein or a protein with at least 80% homology thereto, and said isolated protein lacking its signal peptide or a protein with at least 80% homology thereto. Method for detecting PR07170, PR0361 or PR0846 through interaction with PR01181 and vice versa, method of linking a bioactive molecule to a cell expressing PR01181 using PR07170, PR0361 or PR0846 as binding ligands and vice versa, and method of modulating the activity of PR01181 by binding with PR07170, PR0361 or PR0846, and vice versa, whereby invention 146 is limited to PR07170, invention 147 is limited to PR0361, and invention 148 is limited to PR0846.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/08439

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9839448	A	11-09-1998	AU 6545398 A	22-09-1998
			EP 0972029 A	19-01-2000
			EP 0972030 A	19-01-2000
			WO 9839446 A	11-09-1998
WO 9842741	A	01-10-1998	AU 6777298 A	20-10-1998
			EP 0998490 A	10-05-2000
EP 0834563	A	08-04-1998	JP 10179178 A	07-07-1998
			US 5824504 A	20-10-1998
WO 9707198	A	27-02-1997	US 5707829 A	13-01-1998
			AU 6712396 A	18-02-1997
			AU 6768596 A	12-03-1997
			CA 2227220 A	06-02-1997
			CA 2229208 A	27-02-1997
			EP 0839196 A	06-05-1998
			EP 0851875 A	08-07-1998
			JP 11510045 T	07-09-1999
			US 6043344 A	28-03-2000
			WO 9704097 A	06-02-1997
			US 6074849 A	13-06-2000
			US 5969093 A	19-10-1999
WO 9963088	A	09-12-1999	AU 4328699 A	20-12-1999
			AU 2212299 A	26-07-1999
			WO 9935170 A	15-07-1999
WO 9946375	A	16-09-1999	DE 19811194 A	16-09-1999
WO 9953040	A	21-10-1999	DE 19817557 A	21-10-1999
WO 0008145	A	17-02-2000	AU 5730099 A	28-02-2000
WO 9963083	A	09-12-1999	JP 11332571 A	07-12-1999
			AU 3955099 A	20-12-1999